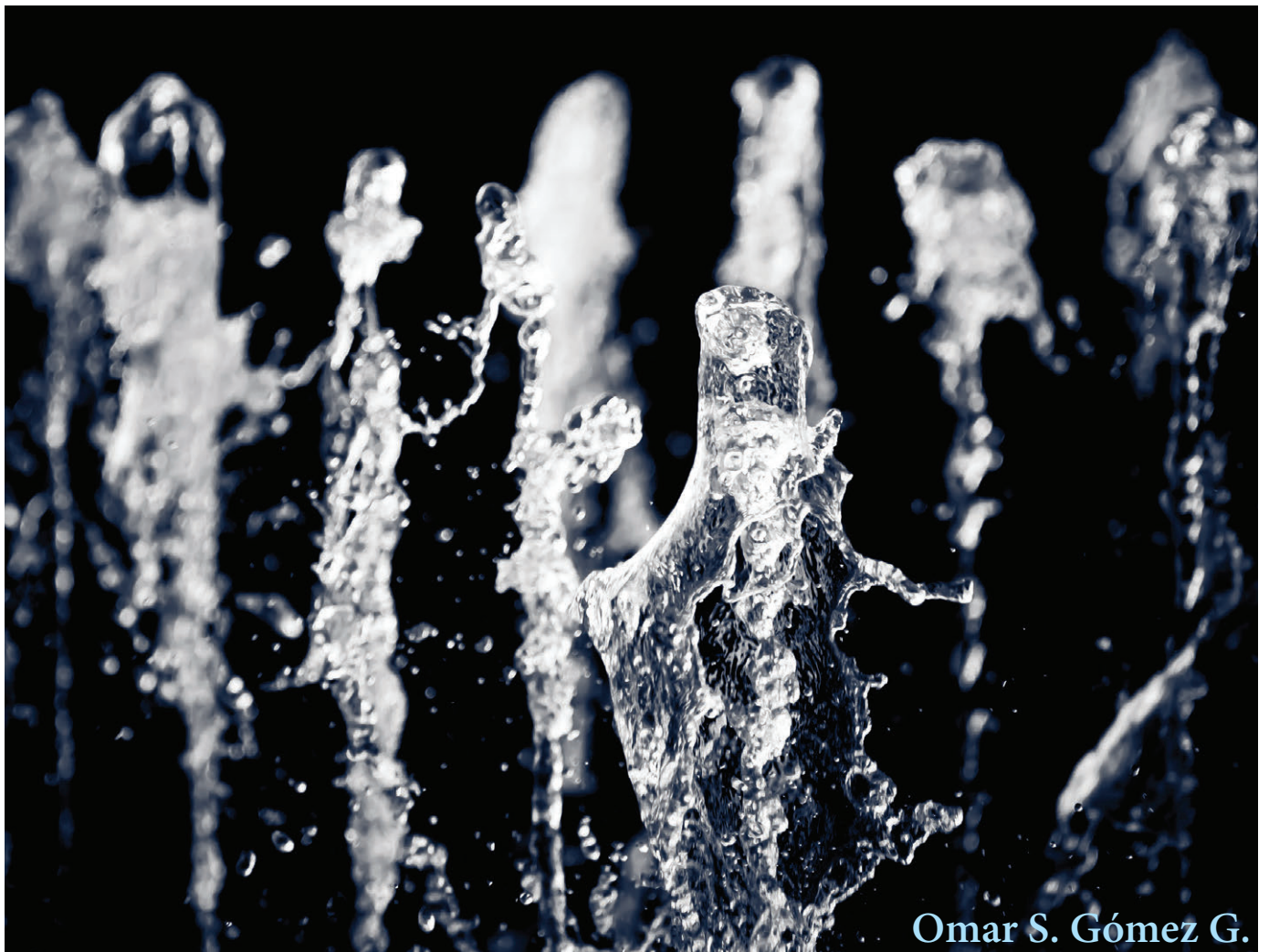


Universidad Politécnica de Madrid

Facultad de Informática

Tipología de Replicaciones para la Síntesis de Experimentos en Ingeniería del Software

TESIS DOCTORAL



Omar S. Gómez G.

2012

Departamento de Lenguajes y Sistemas
Informáticos e Ingeniería de Software
Facultad de Informática
Universidad Politécnica de Madrid



POLITÉCNICA

Tipología de Replicaciones para la Síntesis de
Experimentos en Ingeniería del Software

Tesis que presenta para optar al título de doctor en software y sistemas
Mtro. en Ing. de Software, Omar Salvador Gómez Gómez

Dirigida por
Dra. Natalia Juristo y Dra. Sira Vegas

2012

Documento escrito en \LaTeX utilizando como referencia el proyecto \TeX IS v.1.0+.

Tribunal de Tesis

Dr. Mario G. Piattini

Catedrático de la Universidad de Castilla–La Mancha

(Presidente)

Dr. Óscar Dieste

Profesor de la Universidad Politécnica de Madrid

(Secretario)

Dr. Óscar Pastor

Catedrático de la Universidad Politécnica de Valencia

(Vocal)

Dr. José J. Dolado

Catedrático de la Universidad del País Vasco

(Vocal)

Dr. Andreas Jedlitschka

**Senior Researcher at Fraunhofer Institute for Experimental Software
Engineering**

(Vocal)

*A mis padres,
Ma. del Carmen
y
Salvador
así como a mi esposa
Mayra y mi hija Mayrin*

Agradecimientos

La presente tesis es el resultado de aquella idea vaga que surgió a mediados del año 2005 acerca de indagar y conocer más sobre la Ingeniería del Software Empírica, particularmente, sobre la Ingeniería del Software Experimental.

Este resultado no hubiese sido posible sin la ayuda constante y desinteresada que me ha brindado la profesora Natalia Juristo durante el desarrollo de esta investigación. Gracias profesora Juristo por enseñarme y orientarme en este sinuoso pero gratificante camino de la ciencia y la investigación.

Agradezco también a la profesora Sira Vegas por el apoyo dado en el transcurso de la investigación. Al profesor Oscar Dieste, por estar siempre en la mejor disposición de aclarar las dudas en aspectos específicos de la investigación. A mis compañeros de la facultad Francy, Nella, Anna, Martin, Edison y Rodrigo por alentarme en esta travesía.

De manera especial agradezco al CONACyT (Consejo Nacional de Ciencia y Tecnología) de México por financiar los recursos para llevar a cabo este proyecto. A Georgina Méndez, Verónica Barrientos y Samuel Manterola, por su notable labor que desempeñan en esta institución.

A los profesores Gerardo González[†], Fco. Gerardo Cuellar, Joaquín Mendoza, Francisco Salcedo, Cuauhtémoc Lemus, Carlos Montes de Oca, Miguel Serrano y Gerardo Padilla que directa o indirectamente me han ayudado en las distintas etapas de mi formación.

Son muchas las personas especiales a las que me gustaría agradecer su apoyo, ánimo y compañía en todo este tiempo. A mi padre, Salvador que con amor y sacrificio ha invertido en la medida de sus posibilidades en mi educación. A mi madre, Ma. del Carmen por su amor y apoyo incondicional que siempre me ha dado. Los quiero. También agradezco a mis familiares y amigos por sus palabras de aliento.

A mi esposa Mayra, agradezco su amor, apoyo y compañía que me ha dado en todo momento, gracias por acompañarme en este andar. A mi hija Mayrin, por las alegrías que me ha dado y por estar a mi lado en parte de la redacción de esta tesis. A las dos, las quiero.

Finalmente, a Don Francisco Alonso, agradezco su amistad, apoyo y consejos que me ha dado durante esta estadía. A través de sus charlas e historias he aprendido mucho sobre la cultura de ésta, su tierra. Gracias amigo.

Resumen

La Ingeniería del Software Experimental (ISE) traslada a la Ingeniería del Software (IS) el paradigma experimental que se ha aplicado con éxito en diversas disciplinas científicas. El objetivo de la ISE es hacer de la construcción del software una actividad predecible gracias al conocimiento de las relaciones entre los procesos de producción del software y los productos que se obtienen.

Para avanzar en el paradigma experimental en IS no es suficiente aplicar las técnicas de diseño experimental y el análisis estadístico de datos, sino que es necesario construir una metodología (bien desde cero o adaptada de otras disciplinas) basada en los principios generales del experimentalismo. La motivación principal de esta investigación es trabajar en la adaptación de un aspecto particular del paradigma experimental a la experimentación en IS: la *replicación*.

En ISE se han realizado varias replications de experimentos, sin embargo, aún existe discusión sobre el modo más adecuado de llevarlas a cabo. Algunas preguntas que surgen de esta discusión son: ¿se deben reutilizar los materiales del experimento base?, ¿la replicación debe realizarse de forma independiente, o puede existir algún tipo de comunicación entre experimentadores y replicadores?, ¿qué elementos de la estructura del experimento a replicar pueden variarse y aún considerarse una replicación?

En esta investigación se estudia el concepto de replicación desde una perspectiva teórico-práctica para su incorporación a la ISE. En concreto, se persiguen los siguientes objetivos: 1) estudio del concepto de replicación en distintas disciplinas científicas para tener mayor comprensión de su importación a la ISE, 2) desarrollo de una tipología de replications que ayude a comprender tanto los diferentes tipos de replicación que pueden llevarse a cabo en ISE, así como el papel que cada uno de estos tipos desempeña en la verificación de resultados experimentales y 3) desarrollo de un marco conceptual con ideas clave para comparar conjuntos de replications y obtener conocimiento de ellas que sea de utilidad tanto para el profesional como para el investigador. Para la evaluación de las propuestas de esta tesis se usa un conjunto de 20 replications de diversos autores donde entre otros aspectos se evalúa la efectividad de tres técnicas de evaluación de software.

Abstract

Experimental Software Engineering (ESE) imports the experimental paradigm, successfully applied in many branches of science, to software engineering (SE). ESE aims to make software construction a predictable activity by exploring the relations between software production processes and products.

Advancement of the experimental paradigm in SE entails more than just applying experimental design and statistical data analysis, it also calls for the development of a methodology (either from scratch or adapted from other disciplines) based on the general principles of experimentalism. The primary motivation of this research is to adapt a particular aspect of the experimental paradigm –*replication*– to SE experimentation.

ESE researchers have conducted several replications of experiments. But, there is still debate about what is the best way to go about replication. Some questions raised in this discussion are: Should the baseline experiment materials be reused? Should the replication be carried out independently or is form of communication among experimenters and replicators acceptable? What elements of the experiment structure can be changed for the experiment to still be considered a replication?

This research examines the concept of replication for adoption in ESE from a theoretical and practical viewpoint. It addresses the following objectives: 1) study the concept of replication in different scientific disciplines to gain a better understanding of its importation into ESE; 2) develop a typology of replications useful for understanding what different types of replication can be conducted in ESE and what experimental results verification purpose each replication type serves, and 3) develop a conceptual framework with key ideas for comparing and gathering knowledge about sets of replications that is useful for both practitioners and researchers. To evaluate the proposals of this thesis, we use a set of 20 replications by different authors which examine among other aspects the effectiveness of three software evaluation techniques.

Índice abreviado

I	Planteamiento	1
1	Introducción	3
2	Estado de la Cuestión	9
3	Planteamiento del Problema	21
II	Solución	25
4	Clarificación de Conceptos de Replicación para la ISE	27
5	Tipología de Replicaciones para la ISE	39
6	Marco Conceptual para la Comparación de Replicaciones en ISE	45
7	Procedimiento para Comparar Replicaciones en ISE	67
8	Sub-proceso Generador de Particiones	71
9	Sub-proceso Comparador de Replicaciones	77
III	Evaluación	107
10	Evaluación de Particiones	109
11	Evaluación con Replicaciones Literales	121
12	Evaluación con Replicaciones Diferenciadas-Nativas-Internas. Métrica Faltas Observadas	137
13	Evaluación con Replicaciones Diferenciadas-Nativas-Internas. Métrica Faltas Observables	173
14	Evaluación con Replicaciones Diferenciadas-Conjuntas-Externas	205
15	Evaluación con Replicaciones Diferenciadas-Ajenas-Externas	237

IV Conclusiones	261
16 Conclusiones y Futuras Líneas de Investigación	263
V Apéndices	271
A Descriptions of the Replications Typologies	273
B Técnicas y Herramientas Usadas en la Solución	285
C Códigos empleados en el Comparador de Replicaciones	327
D Formularios del Procedimiento para Comparar Replicaciones	357
E Técnicas de Evaluación de Software	379
F Materiales Empleados en Replicaciones UPM06..10	383
G Tablas con Información de las Replicaciones usadas en la Evaluación	447
Bibliografía	481
Lista de acrónimos	492

Índice extendido

I	Planteamiento	1
1	Introducción	3
1.1	Área de Investigación	3
1.2	Problema	5
1.3	Aproximación a la Solución	6
1.4	Estructura del Documento	7
2	Estado de la Cuestión	9
2.1	Tipos de Replicaciones en ISE	9
2.2	Tipos de Replicaciones en otras Disciplinas	12
2.2.1	Primera Búsqueda de Replicaciones	12
2.2.2	Segunda Búsqueda de Replicaciones	13
2.3	Resultados de ambas Búsquedas	17
3	Planteamiento del Problema	21
3.1	Descripción del Problema	21
3.2	Hipótesis de Trabajo	23
3.3	Estructura de la Solución Propuesta	24
II	Solución	25
4	Clarificación de Conceptos de Replicación para la ISE	27
4.1	Repetición, Replicación y Reproducción	27
4.2	La Repetición como forma de Replicación en ISE	29
4.3	Replicaciones Independientes	32
4.4	Elementos de una Situación Experimental	32
4.4.1	Experimentadores	33
4.4.2	Sitio	33
4.4.3	Protocolo experimental	33
4.4.4	Operacionalizaciones de los Constructos	34
4.4.5	Propiedades de las Poblaciones	35
4.5	Funciones de la Replicación en ISE	36
5	Tipología de Replicaciones para la ISE	39
5.1	Tipología de Replicaciones para la ISE	39

5.1.1	Por Semejanza	39
5.1.2	Por Conexión entre Experimentadores	40
5.1.3	Por Lugar	41
5.2	Cambios Sistemáticos en las Replicaciones	41
6	Marco Conceptual para la Comparación de Replicaciones en ISE	45
6.1	Unidades Comparables entre Replicaciones	46
6.2	Cambios y Nivel de Similitud entre Situaciones Experimentales	50
6.3	Variabilidad de los Tratamientos	54
6.4	Grado de Verificación Logrado	58
6.5	Identificación de Posibles Causas de Resultados no Reproducibles	61
6.6	Comparando Replicaciones	62
7	Procedimiento para Comparar Replicaciones en ISE	67
8	Sub-proceso Generador de Particiones	71
8.1	Identificar Constructos	71
8.2	Generar Tabla de Constructos	72
8.3	Identificar Operacionalizaciones	72
8.4	Generar Tabla de Operacionalizaciones	73
8.5	Generar Tabla de Agrupaciones	73
8.6	Generar Particiones	73
8.7	Seleccionar Particiones	76
9	Sub-proceso Comparador de Replicaciones	77
9.1	Estructura del Comparador de Replicaciones	77
9.2	Generar Tabla Maestra	78
9.3	Sub-proceso Identificación de Similitudes	80
9.3.1	Generar Tablas de Particiones	80
9.3.2	Identificar Similitudes	82
9.4	Sub-proceso Análisis de Resultados Experimentales	84
9.4.1	Generar Tabla de Resultados Experimentales	84
9.4.2	Comparar Resultados Experimentales	85
9.4.3	Generar Tabla con Información del Análisis Realizado	97
9.5	Sub-proceso Empaquetado de Piezas de Conocimiento	99
9.5.1	Generar Tabla de Variaciones en Situación Experimental	100
9.5.2	Generar Pieza de Conocimiento para Profesional	102
9.5.3	Generar Pieza de Conocimiento para Investigador	103
III	Evaluación	107
10	Evaluación de Particiones	109
10.1	Conjunto de Replicaciones Estudiado	109
10.2	Evaluación del Sub-proceso Generador de Particiones	111
10.2.1	Identificar Constructos	111
10.2.2	Generar Tabla de Constructos	112

10.2.3	Identificar Operacionalizaciones	112
10.2.4	Generar Tabla de Operacionalizaciones	113
10.2.5	Generar Tabla de Agrupaciones	115
10.2.6	Generar Particiones	115
10.2.7	Seleccionar Particiones	119
10.3	Resultados de la Evaluación	120
11	Evaluación con Replicaciones Literales	121
11.1	Generar Tabla Maestra	121
11.2	Sub-proceso Identificación de Similitudes	124
11.2.1	Generar Tablas de Particiones	124
11.2.2	Identificar Similitudes	124
11.3	Sub-proceso Análisis de Resultados Experimentales	126
11.3.1	Generar Tabla de Resultados Experimentales	126
11.3.2	Realizar Comparación de Resultados Experimentales	128
11.3.3	Generar Tabla con Información del Análisis Realizado	131
11.4	Sub-proceso Empaquetado de Piezas de Conocimiento	132
11.4.1	Generar Tabla de Variaciones en Situación Experimental	132
11.4.2	Generar Pieza de Conocimiento para Profesional	132
11.4.3	Generar Pieza de Conocimiento para Investigador	133
11.5	Resultados de la Evaluación	135
12	Evaluación con Replicaciones Diferenciadas-Nativas-Internas. Métrica Faltas Observadas	137
12.1	Sub-proceso Identificación de Similitudes	137
12.1.1	Generar Tablas de Particiones	138
12.1.2	Identificar Similitudes	138
12.2	Sub-proceso Análisis de Resultados Experimentales	138
12.2.1	Generar Tabla de Resultados Experimentales	138
12.2.2	Realizar Comparación de Resultados Experimentales	141
12.2.3	Generar Tabla con Información del Análisis Realizado	163
12.3	Sub-proceso Empaquetado de Piezas de Conocimiento	163
12.3.1	Generar tabla de variaciones	164
12.3.2	Generar Pieza de Conocimiento para Profesional	165
12.3.3	Generar Pieza de Conocimiento para Investigador	167
12.4	Resultados de la Evaluación	170
13	Evaluación con Replicaciones Diferenciadas-Nativas-Internas. Métrica Faltas Observables	173
13.1	Sub-proceso Identificación de Similitudes	174
13.1.1	Generar Tablas de Particiones	174
13.1.2	Identificar Similitudes	174
13.2	Sub-proceso Análisis de Resultados Experimentales	174
13.2.1	Generar Tabla de Resultados Experimentales	177
13.2.2	Realizar Comparación de Resultados Experimentales	177
13.2.3	Generar Tabla con Información del Análisis Realizado	195

13.3	Sub-proceso Empaquetado de Piezas de Conocimiento	196
13.3.1	Generar Tabla de Variaciones	197
13.3.2	Generar Pieza de Conocimiento para Profesional	198
13.3.3	Generar Pieza de Conocimiento para Investigador	200
13.4	Resultados de la Evaluación	203
14	Evaluación con Replicaciones Diferenciadas-Conjuntas-Externas	205
14.1	Sub-proceso Identificación de Similitudes	205
14.1.1	Generar Tablas de Particiones	206
14.1.2	Identificar Similitudes	206
14.2	Sub-proceso Análisis de Resultados Experimentales	206
14.2.1	Generar Tabla de Resultados Experimentales	206
14.2.2	Realizar Comparación de Resultados Experimentales	209
14.2.3	Generar Tabla con Información del Análisis Realizado	228
14.3	Sub-proceso Empaquetado de Piezas de Conocimiento	228
14.3.1	Generar Tabla de Variaciones	228
14.3.2	Generar Pieza de Conocimiento para Profesional	230
14.3.3	Generar Pieza de Conocimiento para Investigador	232
14.4	Resultados de la Evaluación	235
15	Evaluación con Replicaciones Diferenciadas-Ajenas-Externas	237
15.1	Sub-proceso Identificación de Similitudes	237
15.1.1	Generar Tablas de Particiones	238
15.1.2	Identificar Similitudes	238
15.2	Sub-proceso Análisis de Resultados Experimentales	238
15.2.1	Generar Tabla de Resultados Experimentales	241
15.2.2	Realizar Comparación de Resultados Experimentales	241
15.2.3	Generar Tabla con Información del Análisis Realizado	254
15.3	Sub-proceso Empaquetado de Piezas de Conocimiento	254
15.3.1	Generar Tabla de Variaciones	255
15.3.2	Generar Pieza de Conocimiento para Profesional	255
15.3.3	Generar Pieza de Conocimiento para Investigador	256
15.4	Resultados de la Evaluación	259
IV	Conclusiones	261
16	Conclusiones y Futuras Líneas de Investigación	263
16.1	Contribuciones de la Investigación	264
16.2	Publicaciones	267
16.3	Futuras Líneas de Investigación	268
V	Apéndices	271
A	Descriptions of the Replications Typologies	273
A.1	Adams et al.	273

A.2	Bahr et al.	273
A.3	Barker and Gurman	273
A.4	Beck	274
A.5	Blomquist	274
A.6	Brown and Coney	274
A.7	Easley et al.	275
A.8	Evanschitzky and Armstrong	275
A.9	Finifter	276
A.10	Finifter	276
A.11	Fuess	276
A.12	Hendrick	276
A.13	Hunter	277
A.14	Hyman and Wright	277
A.15	Van IJzendoorn	278
A.16	Kantowitz et al.	278
A.17	Kelly et al.	278
A.18	La Sorte	279
A.19	Leone and Schultz	279
A.20	Lindsay and Ehrenberg	280
A.21	Lüdtke	280
A.22	Lykken	280
A.23	Mittelstaedt and Zorn	281
A.24	Monroe	281
A.25	Radder	281
A.26	Schmidt	282
A.27	Sidman	282
A.28	Tsang and Kwan	282
B	Técnicas y Herramientas Usadas en la Solución	285
B.1	Mapas Estructurados	285
B.2	Medidas de Proximidad	287
B.3	Cálculo de la Matriz de Similitudes	290
B.4	Análisis de Conglomerados Jerárquico	296
B.5	Análisis de Escalado Multidimensional	297
B.6	Automatización de Cálculos para la Identificación de Similitudes	300
	B.6.1 Generación de Tablas de Contingencia	300
	B.6.2 Cálculo de matriz de distancias	303
	B.6.3 Conglomerados Jerárquico y Escalado Multidimensional	305
B.7	Gráficos de Control	305
B.8	Meta-análisis	310
	B.8.1 Estimación de Tamaños de Efecto Individuales	310
	B.8.2 Estimación de Tamaño de Efecto Global	315
	B.8.3 Heterogeneidad	320
B.9	Automatización de Cálculos para el Análisis de Resultados	321
B.10	Análisis de Causas	322

C	Códigos empleados en el Comparador de Replicaciones	327
C.1	Código fuente escrito en Java del programa <i>Transpose</i>	327
C.2	Fichero <i>r_script01.txt</i> escrito en el lenguaje R	351
C.3	Fichero <i>r_script02.txt</i> escrito en el lenguaje R	355
D	Formularios del Procedimiento para Comparar Replicaciones	357
D.1	Formulario de Constructos de Causa y Efecto (PCR-1)	358
D.2	Formulario de Operacionalizaciones (PCR-2)	359
D.3	Formulario de Agrupaciones (PCR-3)	360
D.4	Formulario de Registro de Situación Experimental (PCR-4)	361
D.5	Formulario de Registro de Tabla Maestra (PCR-5)	364
D.6	Formulario de Registro para Tablas de Particiones (PCR-6)	368
D.7	Formulario de Registro para la Tabla de Resultados (PCR-7)	372
D.8	Formulario de Registro de Posibles Causantes (PCR-8)	373
D.9	Formulario de Registro de Análisis de Resultados (PCR-9)	374
D.10	Formulario de Registro de Variaciones (PCR-10)	375
D.11	Formulario de Registro de Pieza de Conocimiento, Profesional (PCR-11) .	376
D.12	Formulario de Registro de Pieza de Conocimiento, Investigador (PCR-12)	377
E	Técnicas de Evaluación de Software	379
E.1	Lectura de Código por Abstracciones Sucesivas	379
E.2	Pruebas de Caja Blanca o Estructurales	381
E.3	Pruebas de Caja Negra o Funcionales	381
F	Materiales Empleados en Replicaciones UPM06..10	383
F.1	Instrucciones para aplicar la técnica	385
F.2	Formulario de Recogida de datos	387
F.3	Formulario de abstracciones	388
F.4	Formulario de inconsistencias	389
F.5	Formulario de Conclusiones	390
F.6	Instrucciones para aplicar la técnica	391
F.7	Formulario de Recogida de datos	392
F.8	Formulario de Clases de Equivalencia	393
F.9	Formulario de Datos de Prueba	394
F.10	Formulario de Salidas	395
F.11	Formulario de fallos	396
F.12	Formulario de Conclusiones	397
F.13	Instrucciones para aplicar la técnica	398
F.14	Formulario de Recogida de Datos	400
F.15	Formulario de Datos de Prueba	401
F.16	Formulario de Salidas	402
F.17	Formulario de fallos	403
F.18	Formulario de Conclusiones	404
F.19	Código Fuente <i>cmdline</i> (E00)	405
F.20	Especificación del programa “cmdline” (E01)	412
F.21	Suplemento “Revisión de código” y “cmdline” (E02)	414

F.22 Suplemento “Pruebas funcionales” y “cmdline” (E03)	415
F.23 Suplemento “Pruebas estructurales” y “cmdline” (E04)	416
F.24 Código Fuente <i>nametbl</i> (E00)	418
F.25 Especificación del programa “nametbl” (E01)	425
F.26 Suplemento “Revisión de código” y “nametbl” (E02)	427
F.27 Suplemento “Pruebas funcionales” y “nametbl” (E03)	429
F.28 Suplemento “Pruebas estructurales” y “nametbl” (E04)	431
F.29 Código Fuente <i>ntree</i> (E00)	434
F.30 Especificación del programa ‘ntree’ (E01)	439
F.31 Suplemento “Revisión de código” y “ntree” (E02)	441
F.32 Suplemento “Pruebas funcionales” y “ntree” (E03)	443
F.33 Suplemento “Pruebas estructurales” y “ntree” (E04)	444
 G Tablas con Información de las Replicaciones usadas en la Evaluación	 447
 Bibliografía	 481
 Lista de acrónimos	 492

Índice de figuras

4.1	Muestra necesaria para un tamaño de efecto determinado	30
5.1	Cambios sistemáticos para realizar replicaciones	43
6.1	Relación entre constructos y operacionalizaciones	46
6.2	Ejemplo de unidades comparables en varias replicaciones	48
6.3	Dos particiones compuestas de varias unidades comparables	49
6.4	Dos tipos de gráfico que representan la similitud entre replicaciones	54
6.5	Ejemplo de agregación de resultados de nueve replicaciones	55
6.6	Tamaños de efecto y valores de variables de respuesta	56
6.7	Gráfico de dos dimensiones con la efectividad de la técnica A	56
6.8	Ejemplo de variabilidad de la técnica A en seis replicaciones	57
6.9	Adaptación de técnica para análisis de causas	63
6.10	Secuencia de ideas del marco propuesto	64
7.1	Diagrama general del procedimiento para comparar replicaciones	69
7.2	Sub-procesos que conforman este procedimiento	70
8.1	Actividades del sub-proceso <i>generador de particiones</i>	72
8.2	Particiones representadas como un mapa estructurado	75
9.1	Sub-procesos e interacción del <i>comparador de replicaciones</i>	79
9.2	Actividades del sub-proceso <i>identificación de similitudes</i>	80
9.3	Gráficos que representan la similitud entre replicaciones	84
9.4	Actividades del sub-proceso <i>análisis de resultados experimentales</i>	85
9.5	Gráficos de control de dos tratamientos	87
9.6	Ejemplo de meta-análisis	88
9.7	Tipos de agrupaciones de resultados	90
9.8	Distribución de tamaños de efecto estimados	90
9.9	Ejemplos de agrupaciones de resultados consistentes	91
9.10	Gráficos de control de dos tratamientos	92
9.11	Meta-análisis de la partición seleccionada	93
9.12	Representación de similitudes resultante	94
9.13	Gráficos de control tras apartar la replicación $r\mathcal{J}$	95
9.14	Meta-análisis de la partición seleccionada sin replicación $r\mathcal{J}$	96
9.15	Diagrama de causa y efecto con cambios en protocolo	96
9.16	Diagrama de causa y efecto con cambios en poblaciones	97

9.17	Sub-proceso empaquetado de piezas de conocimiento	100
10.1	Particiones resultantes del conjunto de replicaciones	117
10.2	Particiones organizadas de acuerdo a operacionalizaciones	118
11.1	Representación de similitudes, partición 01.a	127
11.2	Gráficos de control resultantes	129
11.3	Meta-análisis de la prueba por PCE y cob. de decisiones	130
12.1	Representación de similitudes, partición 06.b	140
12.2	Gráficos de control resultantes	142
12.3	Meta-análisis de ambas pruebas	143
12.4	Gráficos de control del primer grupo de resultados	145
12.5	Meta-análisis primer grupo de resultados	146
12.6	Gráficos de control del primer grupo de resultados sin upm02	147
12.7	Meta-análisis primer grupo sin upm02	148
12.8	Gráficos de control del primer grupo de resultados sin upm02 y upm03	150
12.9	Meta-análisis primer grupo sin upm02 y upm03	151
12.10	Gráficos de control del segundo grupo de resultados	152
12.11	Meta-análisis segundo grupo	153
12.12	Gráficos de control del segundo grupo de resultados sin upm10	154
12.13	Meta-análisis segundo grupo sin upm10	155
12.14	Cambios identificados en experimentadores y sitio	156
12.15	Cambios identificados en operacionalizaciones	157
12.16	Cambios identificados en experimentadores y sitio	158
12.17	Cambios identificados en el protocolo experimental	158
12.18	Cambios identificados en operacionalizaciones	159
12.19	Cambios identificados en poblaciones	159
12.20	Cambios identificados en protocolo experimental	161
12.21	Cambios identificados en operacionalizaciones	161
12.22	Cambios identificados en poblaciones	162
13.1	Representación de similitudes, partición 11.c	176
13.2	Gráficos de control resultantes	178
13.3	Meta-análisis de ambas pruebas	179
13.4	Gráficos de control del primer grupo de resultados	181
13.5	Meta-análisis del primer grupo de resultados	182
13.6	Gráficos de control del primer grupo de resultados sin upm02	184
13.7	Meta-análisis primer grupo sin upm02	185
13.8	Gráficos de control del segundo grupo de resultados	186
13.9	Meta-análisis del segundo grupo	187
13.10	Gráficos de control del segundo grupo de resultados sin upm08	188
13.11	Meta-análisis segundo grupo sin upm08	189
13.12	Cambios identificados en experimentadores y sitio	190
13.13	Cambios identificados en operacionalizaciones	190
13.14	Cambios identificados en experimentadores y sitio	192

13.15	Cambios identificados en el protocolo experimental	192
13.16	Cambios identificados en operacionalizaciones	193
13.17	Cambios identificados en el protocolo experimental	194
13.18	Cambios identificados en poblaciones	195
14.1	Representación de similitudes, partición 09	208
14.2	Gráficos de control resultantes	210
14.3	Meta-análisis de ambas pruebas	211
14.4	Gráficos de control del primer grupo de resultados	213
14.5	Meta-análisis primer grupo	214
14.6	Gráficos de control del primer grupo de resultados sin ort05	215
14.7	Meta-análisis primer grupo sin ort05	216
14.8	Gráficos de control del primer grupo de resultados	218
14.9	Meta-análisis primer grupo sin ort05, upv05 y uds05	219
14.10	Gráficos de control del segundo grupo de resultados	220
14.11	Meta-análisis segundo grupo	221
14.12	Cambios identificados en experimentadores y sitio	222
14.13	Cambios identificados en el protocolo experimental	222
14.14	Cambios identificados en operacionalizaciones	223
14.15	Cambios identificados en poblaciones	223
14.16	Cambios identificados en el protocolo experimental	226
14.17	Cambios identificados en poblaciones	226
15.1	Representación de similitudes, partición 48	240
15.2	Gráficos de control resultantes	243
15.3	Meta-análisis de las pruebas funcional y estructural	244
15.4	Gráficos de control sin upm02	245
15.5	Meta-análisis de la partición sin upm02	246
15.6	Gráficos de control sin upm02 y upm03	247
15.7	Meta-análisis de la partición sin upm02 y upm03	248
15.8	Gráficos de control resultantes	250
15.9	Meta-análisis de la partición sin umd84, upm02 y upm03	251
15.10	Cambios identificados en experimentadores y sitio	252
15.11	Cambios identificados en el protocolo experimental	252
15.12	Cambios identificados en operacionalizaciones	253
15.13	Cambios identificados en poblaciones	253
B.1	Ejemplo de visualización de particiones	286
B.2	Estructura del fichero que contiene la información de las particiones	287
B.3	Pantalla principal de la herramienta TreeMap	288
B.4	Dendrograma resultante tras aplicar el algoritmo	298
B.5	Gráfico resultante del MDS	300
B.6	Fichero con información de la tabla de partición	301
B.7	Ficheros de texto generados por el programa Transpose	302
B.8	Fichero con tabla de contingencia (sub)elemento 3.1.1	303
B.9	Carga del fichero en el entorno R	304

B.10	Matriz de distancias resultante	304
B.11	Representaciones gráficas de similitudes	306
B.12	Gráfico de control con sus tres zonas	307
B.13	Gráficos de control resultantes	311
B.14	Diagrama de bosque	319
B.15	Fichero de texto con tabla de resultados experimentales	322
B.16	Gráficos de control resultantes	323
B.17	Meta-análisis de la partición seleccionada	324
B.18	Diagrama de Ishikawa usado en el análisis de causas	325
B.19	Diagrama de Ishikawa para el protocolo experimental	325
B.20	Diagrama de Ishikawa con número de cambios identificados	326

Índice de tablas

2.1	Tipos de replicación y la relación identificada entre estos tipos	11
2.2	Clasificaciones de replicaciones por área	12
2.3	Clasificaciones más citadas	13
2.4	Número de documentos resultantes	14
2.5	Citas de las clasificaciones identificadas	15
2.6	Cadenas de búsqueda utilizadas con el término <i>reproduction</i>	16
2.7	Clasificaciones encontradas en las dos búsquedas	17
2.8	Grupos de replicaciones identificados	19
2.9	Elementos identificados que pueden variar en la replicación	20
4.1	Conocimiento ganado en la replicación	36
6.1	Registro de cambios del protocolo experimental en tres replicaciones . . .	51
6.2	Conocimiento ganado en la replicación	52
6.3	Ejemplo de similitudes representadas como distancias	53
6.4	Situaciones presentadas en una partición	60
6.5	Ejemplo de conocimiento obtenido	61
8.1	Ejemplo de tabla con registro de constructos	72
8.2	Ejemplo de tabla con registros de operacionalizaciones	73
8.3	Ejemplo de pares de constructos y operacionalizaciones identificados . . .	74
8.4	Ejemplo de particiones resultantes en cuatro replicaciones	74
8.5	Partición válida para usar en el <i>comparador de replicaciones</i>	76
9.1	Fragmento de la tabla maestra	81
9.2	Sección de la tabla de partición referente al protocolo experimental . . .	82
9.3	Matriz de similitud de una partición con tres replicaciones	83
9.4	Resultados experimentales de tres replicaciones	85
9.5	Combinaciones de resultados que pueden obtenerse	88
9.6	Cambios identificados	98
9.7	Información de las comparaciones	100
9.8	Tabla de variaciones en situación experimental	101
9.9	Tabla de partición referente a experimentadores y sitio	102
9.10	Pieza de conocimiento para profesional	104
9.11	Pieza de conocimiento para investigador	106
10.1	Replicaciones usadas para la evaluación	110

10.2	Tabla de constructos del formulario PCR-1	112
10.3	Tabla de operacionalizaciones del formulario PCR-2	114
10.4	Tabla de agrupaciones del formulario PCR-3	115
10.5	Particiones identificadas en el conjunto de replicaciones	116
10.6	Conjunto inicial de particiones seleccionadas	119
10.7	Particiones a utilizar	119
11.1	Elementos identificados en experimentadores y sitio	123
11.2	Fragmento de tabla de particiones	125
11.3	Resultados experimentales de la partición 01.a	126
11.4	Información de las comparaciones, partición 01.a	131
11.5	Tabla de variaciones en situación experimental	132
11.6	Pieza de conocimiento para profesional	133
11.7	Pieza de conocimiento para investigador	134
12.1	Tabla de particiones referente a los experimentadores y sitio	139
12.2	Resultados experimentales de la partición 06.b	141
12.3	Cambios identificados en primer grupo	156
12.4	Cambios identificados en segundo grupo	160
12.5	Cambios identificados en inter-grupos	161
12.6	Información de las comparaciones primer grupo	163
12.7	Información de las comparaciones segundo grupo	164
12.8	Tabla de variaciones primer grupo	165
12.9	Tabla de variaciones segundo grupo	166
12.10	Pieza de conocimiento, profesional primer grupo	167
12.11	Pieza de conocimiento, profesional segundo grupo	168
12.12	Pieza de conocimiento para investigador del primer grupo de resultados .	169
12.13	Pieza de conocimiento para investigador del segundo grupo de resultados	170
13.1	Tabla de particiones referente a los experimentadores y sitio	175
13.2	Resultados experimentales de la partición 11.c	177
13.3	Cambios identificados en primer grupo	191
13.4	Cambios identificados en segundo grupo	193
13.5	Cambios identificados en inter-grupos	194
13.6	Información de las comparaciones primer grupo	196
13.7	Información de las comparaciones segundo grupo	197
13.8	Tabla de variaciones primer grupo	198
13.9	Tabla de variaciones segundo grupo	199
13.10	Pieza de conocimiento, profesional primer grupo	200
13.11	Pieza de conocimiento, profesional segundo grupo	201
13.12	Pieza de conocimiento para investigador del primer grupo de resultados .	202
13.13	Pieza de conocimiento para investigador del segundo grupo de resultados	203
14.1	Tabla de particiones referente a los experimentadores y sitio	207
14.2	Resultados experimentales de la partición 09	209
14.3	Cambios identificados en primer grupo	224

14.4	Cambios identificados en inter-grupos	227
14.5	Información de las comparaciones primer grupo	228
14.6	Información de las comparaciones segundo grupo	229
14.7	Tabla de variaciones primer grupo	230
14.8	Tabla de variaciones segundo grupo	231
14.9	Pieza de conocimiento, profesional primer grupo	232
14.10	Pieza de conocimiento, profesional segundo grupo	233
14.11	Pieza de conocimiento para investigador del primer grupo de resultados .	234
14.12	Pieza de conocimiento para investigador del segundo grupo de resultados	235
15.1	Tabla de particiones referente a los experimentadores y sitio	239
15.2	Resultados experimentales de la partición 48	241
15.3	Información de las comparaciones de resultados	255
15.4	Tabla de variaciones	256
15.5	Pieza de conocimiento, profesional	257
15.6	Pieza de conocimiento resultante para investigador	258
B.1	Estructura de la tabla de contingencia	289
B.2	Técnicas de análisis de datos usadas en replicaciones A y B	289
B.3	Técnicas de análisis de datos que utilizan dos replicaciones	290
B.4	Tabla de contingencia elemento <i>Técnica de análisis de datos</i>	291
B.5	Matriz de similitudes elemento <i>Técnica de análisis de datos</i>	291
B.6	Matriz de similitudes, experimentadores	292
B.7	Matriz de similitudes, sitio	292
B.8	Matriz de similitudes, protocolo experimental	292
B.9	Matriz de similitudes, operacionalizaciones	292
B.10	Matriz de similitudes, poblaciones	292
B.11	Ejemplo de elementos en experimentadores y sitio	293
B.12	Ejemplo de elementos en protocolo experimental	293
B.13	Ejemplo de elementos en operacionalizaciones de los constructos	294
B.14	Ejemplo de elementos en poblaciones	294
B.15	Matriz de similitudes que representa la situación experimental	295
B.16	Distancias en espacio euclídeo	296
B.17	Matriz de distancias de las replicaciones $r1$, $r2$ y $r3$	297
B.18	Matriz resultante en primera iteración	297
B.19	Coordenadas resultantes de cada replicación	299
B.20	Resultados experimentales de tres replicaciones	313
B.21	Resultados de la replicación $r1$	314
B.22	Tamaños de efecto estimados e intervalos de confianza	315
B.23	Resultados de W , WY , WY^2 y W^2 para calcular Q y C	318
B.24	Resultados de varianza total y peso en replicaciones de una partición . .	318
B.25	Resultados tras aplicar las funciones	319
G.1	Tabla de agrupaciones del formulario PCR-3	448
G.2	Tabla de agrupaciones del del formulario PCR-3 (continuación)	449
G.3	Particiones identificadas en el conjunto de replicaciones	450

G.4	Particiones identificadas en el conjunto de replicaciones (continuación) . .	451
G.5	Tabla maestra, experimentadores y sitio	452
G.6	Tabla maestra, experimentadores y sitio (cont.)	453
G.7	Tabla maestra, protocolo experimental	454
G.8	Tabla maestra, protocolo experimental (cont.)	455
G.9	Tabla maestra, protocolo experimental (cont.)	456
G.10	Tabla maestra, operacionalizaciones	457
G.11	Tabla maestra, operacionalizaciones (cont.)	458
G.12	Tabla maestra, operacionalizaciones (cont.)	459
G.13	Tabla maestra, poblaciones	460
G.14	Tabla maestra, poblaciones (cont.)	461
G.15	Tabla maestra, poblaciones (cont.)	462
G.16	Partición 01.a, experimentadores, sitio y protocolo experimental	463
G.17	Partición 01.a, operacionalizaciones y poblaciones	464
G.18	Partición 06.b, experimentadores y sitio	465
G.19	Partición 06.b, protocolo experimental	466
G.20	Partición 06.b, operacionalizaciones y poblaciones	467
G.21	Partición 11.c, experimentadores y sitio	468
G.22	Partición 11.c, protocolo experimental	469
G.23	Partición 11.c, operacionalizaciones y poblaciones	470
G.24	Partición 09, experimentadores y sitio	471
G.25	Partición 09, protocolo experimental	472
G.26	Partición 09, operacionalizaciones	473
G.27	Partición 09, poblaciones	474
G.28	Partición 48, experimentadores y sitio	475
G.29	Partición 48, experimentadores y sitio (cont.)	476
G.30	Partición 48, protocolo experimental	477
G.31	Partición 48, operacionalizaciones	478
G.32	Partición 48, poblaciones	479

Parte I

Planteamiento

Capítulo 1

Introducción

—Paréceme, Sancho, que no hay refrán que no sea verdadero, porque todos son sentencias sacadas de la misma experiencia, madre de las ciencias todas.

Tercera parte del ingenioso hidalgo
Don Quijote de la Mancha
Miguel de Cervantes

RESUMEN: En este capítulo se describe el área de investigación en que se enmarca esta tesis, el problema al que este trabajo pretende contribuir, los objetivos de investigación y la aproximación utilizada para aportar una solución.

1.1 Área de Investigación

Hoy en día los ingenieros del software cuentan con una variedad de tecnologías como son herramientas, lenguajes y entornos de programación, metodologías, técnicas y procesos que pueden utilizar en las distintas etapas de construcción de sistemas software. A menudo, los ingenieros del software tienen la necesidad de evaluar y seleccionar en base a determinados criterios las tecnologías que mejor se adecuen a la construcción o mantenimiento de un producto software. Realizar una mala evaluación y selección de estas tecnologías puede ocasionar retrasos en la entrega del producto, incrementos en los costes, o defectos en el producto resultante.

Actualmente, la comunidad científica en IS (Ingeniería del Software) está de acuerdo en la carencia de una base teórica sólida para la disciplina (Tichy, 1998; Kitchenham et al., 2002; Hannay et al., 2007; Sjøberg et al., 2008; Shull y Feldmann, 2008). Esto significa que no existe conocimiento suficiente para realizar una evaluación y selección adecuada de la gama de tecnologías que dispone el ingeniero del software.

En IS, los resultados de la aplicación de una determinada tecnología para la construcción de software es, a día de hoy, impredecible (Wohlin et al., 1999). A menudo se adoptan tecnologías nuevas sin contar con evidencia convincente de que serán efectivas (Zelkowitz et al., 2003). No existe evidencia alguna que apoye la mayoría de las creencias en que basamos la construcción de sistemas software (Juristo et al., 2002).

A través de la experimentación es posible evaluar de manera objetiva las distintas tecnologías disponibles para la construcción de sistemas software. Por ejemplo, en el área de pruebas del software, un experimento¹ permite comparar de manera objetiva la efectividad de varias técnicas de detección de faltas. La experimentación contrasta las creencias y las opiniones con la realidad convirtiéndolas en conocimiento científico o desechándolas. En la historia de todas las disciplinas ingenieriles se ha producido una transición desde las creencias, especulaciones y aciertos casuales hasta el conocimiento científico mediante el cual una ingeniería alcanza resultados predecibles (Shaw, 1990, 2009).

A comienzos del Siglo XI Ibn al Haytham (conocido también como Alhacén o Alhazen) (Alhazen, 1021) demostró a través de experimentos que la percepción visual es causada por la luz que se dirige desde su fuente externa tras reflejarse en los objetos que vemos. Esta demostración refutó la teoría existente de Euclides y Ptolomeo según la cual la luz es emitida por el ojo del observador en busca del objeto y hacia éste, así como refutó la teoría de Aristóteles y seguidores que creían que las formas físicas de los objetos entraban hacia los ojos (Hull, 2011; Wikipedia, 2011). Alhazen fue un científico pionero en el uso de la experimentación para contrastar las teorías establecidas en aquel entonces.

Aunque la experimentación tiene sus orígenes en ciencias como la física o la química, en la actualidad diversas disciplinas científicas emplean la experimentación como forma de obtener conocimientos válidos. Estas disciplinas han adaptado el método experimental a sus propios contextos. Por ejemplo, (Campbell y Stanley, 1963), Cook y Campbell (Cook y Campbell, 1976), Cohen (Cohen, 1988), entre otros, han contribuido en la adaptación del método experimental en disciplinas de las ciencias sociales como son la psicología, sociología y antropología. Otro ejemplo se presenta en economía donde los trabajos de Smith (Smith, 1976, 1982)², Friedman y Sunder (Friedman y Sunder, 1994), Guala (Guala, 2005), Croson (Croson, 2005), entre otros, han contribuido a trasladar el paradigma experimental a esta disciplina.

La aplicación del paradigma experimental aplicado a la investigación en IS cuenta ya con varias décadas. Los orígenes de la ISE (Ingeniería del Software Experimental) se remontan a los trabajos de Weinberg y Gressett (Weinberg y Gressett, 1963), Grant y Sackman (Grant y Sackman, 1967), Kolence (Kolence, 1973) y Hanson (Hanson, 1977) donde los autores utilizan aproximaciones científicas basadas en la observación y evaluación de tecnologías para la construcción de sistemas software. No obstante, los primeros experimentos así como la aplicación del paradigma experimental a la IS fue hecha por Basili en los años 80's (Basili et al., 1986). La experimentación en IS hace posible la identificación de variables relevantes en la construcción de software y la comprensión de relaciones existentes entre ellas (Pfleeger, 1999).

Para aplicar el método experimental en una disciplina, es necesario construir una metodología particular basada en los principios generales de la experimentación. Si se importa directamente la metodología de una disciplina a otra se corre el riesgo de que ésta no sea completamente funcional.

En IS se han realizado algunas adaptaciones del paradigma experimental. Por ejemplo, Basili et al. (Basili et al., 1986), Wohlin et al. (Wohlin et al., 1999), Perry et al. (Perry et al., 2000), Lott y Rombach (Lott y Rombach, 1996) así como Juristo y Moreno

¹En este documento, el término experimento se usa como sinónimo de experimento controlado.

²Los trabajos de Smith en el área de economía experimental lo llevaron a obtener el premio nobel en economía en el año 2002. A raíz de su trabajo, la experimentación en esta disciplina comenzó a ser ampliamente aceptada como una metodología válida.

(Juristo y Moreno, 2001) han trabajado en la adaptación de procesos para la realización de experimentos. Por otra parte, Singer (Singer, 1999) propone adaptar las guías de estilo de la APA (*American Psychological Association*) para reportar experimentos en IS. De manera similar, Kitchenham et. al. (Kitchenham et al., 2002) así como Jedlitschka y Pfahl (Jedlitschka y Pfahl, 2005) han desarrollado guías para reportar experimentos controlados.

La motivación principal de la presente investigación es trabajar en la adaptación de un aspecto particular del paradigma experimental a la experimentación en IS que ha rehusado, desde hace años, resistiéndose a un tratamiento en la ISE. Se trata de la repetición o replicación de experimentos.

1.2 Problema

En sus inicios, la ISE se centró en estudiar la aplicación de los principios del laboratorio y el experimento a la IS. Tras un par de décadas, la realización de experimentos en ISE se ha convertido en una tarea bastante bien comprendida. Pero la realización de experimentos aislados es sólo un primer paso en la secuencia de actividades del paradigma experimental. Queda pendiente el análisis y adaptación de otros principios del experimentalismo. Actualmente, la investigación en ISE se centra, no ya en la realización de experimentos de laboratorio a pequeña escala (aislados) sino pasar a la experimentación a gran escala realizando experimentos coordinados (Selby, 2007) y concatenando conjuntos de experimentos (Pfleeger, 2005). La migración de experimentos de laboratorio aislados hacia la combinación de varios experimentos permitirá pasar de resultados locales poco generalizables a evidencias globales sobre la construcción de software.

Para llegar a la generalización de resultados en ISE, es necesario repetir un experimento en distintos contextos; es decir, contar con varias replications. Una vez que se tienen varias replications, debe llevarse a cabo la síntesis de resultados para obtener alguna evidencia concluyente o pieza de conocimiento del evento estudiado.

En ISE, la realización de replications ha sido complicada debido al estado de inmadurez en que todavía se haya el paradigma experimental aplicado a la construcción de software. Dado el desconocimiento de los mecanismos que regulan la construcción de software así como las condiciones experimentales relevantes para su estudio, cualquier cambio en la replicación por mínimo que sea, puede ocasionar alteraciones inexplicables en los resultados. Sin embargo, a menudo las modificaciones al contexto donde se replica el experimento son inevitables. Estas modificaciones pueden ocasionar variaciones importantes en los resultados de la replicación que hacen difícil (por no decir imposible) verificar los resultados del experimento base.

En varios intentos de combinar resultados de replications en ISE (Hayes, 1999; Miller, 1999, 2000, 2005; Runeson et al., 2006; Hannay et al., 2009; Jørgensen, 2004; Pickard et al., 1998; Shull et al., 2002; Juristo et al., 2004), los investigadores han reportado que los resultados de las replications son tan diferentes que resulta complicado agregarlos. Se han introducido tantos cambios en las replications que resulta imposible extraer consecuencias de la comparación de resultados.

Esta inmadurez experimental lleva a una incomprensión del concepto de replicación aplicado a los experimentos en IS. Cuando se piensa en la replicación se tiene como estereotipo aquella que se lleva a cabo independientemente por otros investigadores en sitios diferentes a los del experimento base. Sin embargo este estereotipo no se corresponde

con la mayoría de las replications en ISE que son realizadas o de manera conjunta, o reutilizando parte de los materiales empleados en el experimento base, o en el mismo sitio (Basili y Selby, 1987; Porter et al., 1995; Fusaro et al., 1997; Miller et al., 1998; Sandahl et al., 1998; Porter y Votta, 1998; Wood et al., 1997; Juristo y Vegas, 2003; Vegas et al., 2006; Juristo y Vegas, 2009). Esta situación parece ocasionar confusión sobre si estas replications violan o no el principio de independencia.

El modo de realizar replications ha originado cierta controversia en ISE. Por una parte están quienes recomiendan realizar replications reutilizando parte de los materiales del experimento base (Basili et al., 1999; Shull et al., 2008) con el fin de realizar replications lo más similar posible. Por otra parte, están quienes recomiendan utilizar protocolos y materiales distintos a los del experimento base (Kitchenham, 2008; Miller, 2005) con el fin de preservar el principio de independencia y evitar propagar errores en replications que utilicen los mismos materiales. Algunos incluso sugieren emplear formas alternativas de verificación (Miller, 2008) de los resultados experimentales con el fin de entender los problemas que la replicación ha tenido hasta ahora en los experimentos en IS. Estas controversias tienen su origen en una falta de adaptación adecuada a ISE del principio de la replicación del paradigma experimental.

Puntualizando lo anterior, hoy en día no existe en ISE trabajo que estudie en profundidad el concepto de replicación. Tampoco existen trabajos específicos que pretendan hacer una clasificación o tipología de replications como tema principal de investigación. Por último, aunque en ISE se han usado técnicas para la combinación de resultados experimentales, se es incapaz de extraer consecuencias cuando los resultados se han o no reproducido, no se han desarrollado ideas que permitan explotar conjuntos de replications con el fin de conocer el grado de verificación alcanzado cuando se tienen resultados reproducibles, o identificar variables contextuales críticas en la construcción de software cuando se obtienen resultados no reproducibles.

De acuerdo a la problemática descrita, el **objetivo general** de esta investigación es comprender y adaptar el concepto de replicación a la ISE con el fin de contribuir a clarificar la polémica antes descrita, así como proporcionar a los investigadores guías para realizar distintos tipos de replications útiles, compararlas y obtener conocimiento de ellas.

1.3 Aproximación a la Solución

La solución propuesta parte del estudio del concepto de replicación tanto en la literatura en ISE como en la literatura de otras disciplinas científicas. En estas disciplinas se examina con detalle los diferentes tipos y formas de replicación comúnmente aceptados.

Con los resultados del estudio, se identifican los elementos que conforman un experimento o situación experimental típica y los posibles cambios que puede sufrir. Estos cambios establecen los distintos tipos de replicación que pueden realizarse en ISE.

Por último se desarrolla un marco conceptual con ideas clave para realizar comparaciones entre replications y obtener conocimiento de ellas. Las ideas de este marco se trasladan a un procedimiento que facilita al investigador realizar comparaciones entre replications.

Para la evaluación de la tipología de replications y del marco conceptual se usa un conjunto de veinte replications afines al área de evaluación de software. Estas replications se han realizado en distintos sitios y por distintos investigadores.

1.4 Estructura del Documento

Este documento se estructura en cinco partes que a continuación se describen:

1. Planteamiento. Este apartado contiene la introducción a la investigación realizada y se compone de tres capítulos: introducción (capítulo 1); estado de la cuestión (capítulo 2) y el planteamiento del problema (capítulo 3).
2. Solución. En este apartado se describe la solución propuesta por esta investigación. En el capítulo 4 se presenta la clarificación de conceptos de replicación para la ISE. En el capítulo 5 se describe la tipología de replications propuesta. En el capítulo 6 se presenta el marco conceptual con las ideas clave para realizar comparaciones entre replications. Siguiendo las ideas del marco conceptual, en los restantes tres capítulos (caps. 7, 8 y 9) se describe un procedimiento para llevar a cabo comparaciones entre replications.
3. Evaluación. En este apartado se lleva a cabo la evaluación de la solución propuesta. La evaluación de la tipología y el marco conceptual se realiza a través del procedimiento para comparar replications que se compone de dos sub-procesos. En el capítulo 10 se evalúa el primer sub-proceso con un conjunto de replications afines a la IS, y en los capítulos 11, 12, 13, 14 así como 15 se evalúa el segundo sub-proceso con distintos tipos de replications del conjunto seleccionado.
4. Conclusiones. Este apartado contiene un único capítulo (cap. 16) donde se presentan las conclusiones de la investigación así como futuras líneas de investigación.
5. Apéndices. En este último apartado se incluye la bibliografía utilizada en esta investigación así como algunos apéndices relevantes.

Capítulo 2

Estado de la Cuestión

RESUMEN: En este capítulo se lleva a cabo una revisión de las formas de replicación descritas en algunos trabajos de ISE así como los tipos de replications que han sido propuestos en otras disciplinas científicas.

2.1 Tipos de Replicaciones en ISE

Como punto de partida se realizó una revisión en la literatura de ISE para conocer las formas o tipos de replicación aceptados. No se encontró ningún trabajo específico que pretenda hacer una tipología o clasificación de replications. No obstante se han encontrado tres trabajos que han necesitado elaborar una clasificación de replications como parte de la investigación realizada.

El primer trabajo se localiza en un artículo escrito por Basili et al. (Basili et al., 1999). En donde los autores presentan un marco de trabajo para organizar conjuntos de estudios relacionados. A través del artículo los autores van describiendo los diferentes aspectos del marco de trabajo, en uno de los aspectos los autores definen una clasificación de replications que se compone de tres categorías principales, en dos de estas categorías se definen varios tipos de replications. Los autores ejemplifican la clasificación con diferentes replications que han realizado. En total la clasificación se compone de seis tipos de replications:

- *Replicaciones estrictas.* Son replications lo más exactas posibles al experimento original que incrementan la confianza en la validez de la conclusión del experimento. Estas replications demuestran que los resultados del experimento original son repetibles y que se han reportado de manera exacta por los investigadores del estudio original.
- *Replicaciones que varían la forma en que el experimento es ejecutado.* Incrementan la confianza en los resultados experimentales, probando las mismas hipótesis del experimento original pero alterando los detalles del experimento para que ciertas amenazas a la validez interna pueden ser estudiadas.

- *Replicaciones que varían las variables intrínsecas del objeto de estudio (variables independientes)*. Investigan los aspectos de las variables independientes que son importantes, variando sistemáticamente sus propiedades intrínsecas y examinando los resultados.
- *Replicaciones que varían las variables intrínsecas respecto al enfoque de la evaluación (variables dependientes)*. Varían las formas en cómo la variable dependiente es medida, para comprender en qué dimensiones de la tarea, una variable independiente obtiene la mayor ganancia.
- *Replicaciones que varían las variables de contexto en el entorno donde la solución es evaluada*. Pretenden identificar factores importantes del entorno que afectan los resultados y así entender la validez externa de los mismos.
- *Replicaciones que extienden la teoría*. Estas replicaciones ayudan a determinar los límites donde se mantienen los resultados. Estas replicaciones realizan grandes cambios a las variables independientes, dependientes o al contexto para ver si los principios básicos se mantienen.

La segunda clasificación se encuentra en una tesis de Máster (Almqvist, 2006) que tiene como objetivo estudiar el uso de la replicación de experimentos controlados en Ingeniería del Software Empírica. El autor realiza un sondeo (*survey*) a 44 artículos en los que se describen 51 experimentos controlados y 31 replicaciones. Como método para identificar los artículos relevantes, se realizó una revisión sistemática. En alguna parte del trabajo (capítulo 4), se definen varias categorías para organizar los experimentos encontrados. En una de las categorías se desarrolla una clasificación para categorizar las replicaciones encontradas. Almqvist (Almqvist, 2006) toma como referencia los tipos de replicación descritos en la clasificación de Lindsay y Ehrenberg (Lindsay y Ehrenberg, 1993), y añade los tipos de replicación *interna* y *externa*. Por tanto, define los siguientes cuatro tipos de replicaciones:

- *Replicaciones diferenciadas externas*. Son replicaciones hechas por investigadores distintos a los del experimento original. En estas replicaciones se varía intencionalmente diferentes condiciones del experimento.
- *Replicaciones diferenciadas (mejoradas) internas*. Son hechas por los mismos investigadores bajo distintas condiciones, en distintos entornos, o con tareas modificadas.
- *Replicaciones cercanas (similares) internas*. Los investigadores del experimento original replican el experimento o, la replicación se realiza en la misma institución donde se realizó el experimento base. Las condiciones utilizadas son similares a las del experimento previo.
- *Replicaciones cercanas (similares) externas*. Son replicaciones hechas por investigadores distintos en instituciones distintas a las del experimento original utilizando condiciones similares.

La tercera clasificación es muy reciente (Krein y Knutson, 2010). En un punto del marco unificador para la organización de los métodos de investigación en IS los autores elaboran una taxonomía de replicaciones donde definen cuatro tipos de replicaciones:

- *Replicación estricta*. Aquella donde se replica un estudio previo de forma lo más precisa posible.
- *Replicación diferenciada*. Aquella donde intencionadamente se alteran aspectos del estudio previo para evaluar los límites de sus conclusiones.

- *Replicación dependiente*. Un estudio específicamente diseñado con referencia a uno o más estudios previos que pretende ser un estudio replicado.
- *Replicación independiente*. Sigue las mismas cuestiones y/o hipótesis del estudio previo, pero se realiza sin conocimiento de este estudio previo porque el investigador no es consciente del trabajo previo o porque se desea evitar la contaminación con prejuicios de los investigadores anteriores.

Como se observa, la clasificación de Basili et al. (Basili et al., 1999) contiene descripciones más completas que las dos clasificaciones restantes. De acuerdo a las descripciones de estas replications es posible identificar cierta relación. En la tabla 2.1 se muestran los tipos de replicación de estas clasificaciones así como la relación identificada entre estos tipos.

Basili et al.	Almqvist	Krein y Knutson
Estricta	Cercana interna Cercana externa	Estricta
Varía la forma de ejecución del experimento		
Varía variables independientes	Diferenciada externa	Diferenciada
Varía variables dependientes	Diferenciada interna	
Varía variables de contexto		
Extiende la teoría		
		Dependiente
		Independiente

Tabla 2.1: Tipos de replicación y la relación identificada entre estos tipos

Los autores de estas clasificaciones concuerdan en definir una forma de replicar un experimento lo más similar posible, Basili et al. (Basili et al., 1999) así como Krein y Knutson (Krein y Knutson, 2010) denominan esta forma como *replicación estricta* mientras que Almqvist (Almqvist, 2006) la llama *replicación cercana interna*. Respecto a la replicación diferenciada descrita en (Almqvist, 2006) y (Krein y Knutson, 2010) sólo se menciona que son replications con ciertos cambios, al no proporcionar más información, estas replications pueden relacionarse con varias replications descritas en (Basili et al., 1999) donde también se mencionan cambios en la replicación.

Como el objetivo de estos tres trabajos no es en sí elaborar una clasificación sino usarla para otros intereses, quedan pendientes varias cuestiones por resolver y clarificar. Por ejemplo, en las clasificaciones no se discute de manera detallada los diferentes tipos de cambios que puede haber en la replicación. Esto implica que en la práctica cuando se esté conociendo alguna replicación se dude sobre el tipo de replicación que es.

Otra cuestión importante que no se aborda es el papel que desempeña cada tipo de replicación en la verificación de resultados experimentales. Una vez que se conoce el tipo de replicación a realizar no se sabe cuál es su contribución en la verificación de resultados respecto a los resultados obtenidos con anterioridad, ¿Obedecen los resultados de alguno de estos tipos de replicación a alguna realidad de la construcción de software? o ¿Son estos resultados artificiales producto del contexto donde se realizó la replicación?

Existen otros trabajos en ISE que mencionan tipos de replications pero no realizan una clasificación. En Brooks et al. (Brooks et al., 1996), y Mendonça et al. (Mendonça et al., 2008) se mencionan las diferencias entre la replicación *interna* y *externa*. Shull et al. (Shull et al., 2008) comentan algunos tipos de replications (*exacta*, *independiente*, *dependiente* y *conceptual*) para describir el papel que juega la replicación en IS empírica. Lung et al. (Lung et al., 2008) mencionan dos tipos de replicación (*literal* y *teórica*) para

explicar el tipo de replicación que realizaron. Por último Mandić et al. (Mandić et al., 2009) discuten dos tipos de replicaciones: *exacta* o *parcial* y replicaciones con el objetivo de mejorar el estudio original.

Tal y como se esperaba, se encontraron muy pocos trabajos en ISE sobre el concepto y los tipos de replicación. Se buscó entonces en otras disciplinas científicas los tipos de replicación comúnmente aceptados así como las variaciones aceptadas entre replicaciones.

2.2 Tipos de Replicaciones en otras Disciplinas

Con el fin de conocer más sobre las variaciones aceptadas en una replicación, se han examinado varias clasificaciones de replicaciones usadas en distintas disciplinas científicas. Las clasificaciones encontradas se obtuvieron tras realizar dos búsquedas que a continuación se describen.

2.2.1 Primera Búsqueda de Replicaciones

En esta primera búsqueda se utilizaron los motores Google[®], Google scholar[®], JSTOR[®] y ScienceDirect[®] empleando los términos: *types of replications*, *types of experimental replications*, *typology of replications*, *replication types*, *replication typologies* así como *classification of replications*. Tras realizar las búsquedas en los cuatro motores y examinar los resultados arrojados, se localizó un conjunto inicial de 10 clasificaciones de replicaciones (Van IJzendoorn, 1994; Evanschitzky y Armstrong, 2010; Kantowitz et al., 1984; Bahr et al., 1983; Tsang y Kwan, 1999; Mittelstaedt y Zorn, 1984; Lindsay y Ehrenberg, 1993; Easley et al., 2000; Hunter, 2001; Schmidt, 2009). Este conjunto inicial de clasificaciones sirvió como fuente para localizar más clasificaciones de replicaciones. Siguiendo las referencias de este conjunto inicial fue posible localizar otras 8 clasificaciones (Lykken, 1968; Hendrick, 1990; Finifter, 1972; La Sorte, 1972; Kelly et al., 1979; Leone y Schultz, 1980; Monroe, 1992; Radder, 1992).

Por tanto, se seleccionaron 18 clasificaciones de replicaciones. De manera conjunta las clasificaciones contienen un total de 79 tipos de replicaciones. De estas clasificaciones 11 (61%) pertenecen a áreas de ciencias sociales, 6 (33%) al área de negocios y 1 (6%) al área de filosofía. En el apéndice A se encuentran las descripciones de los distintos tipos de replicaciones encontrados. En la tabla 2.2 se muestran las clasificaciones agrupadas por área y por autores.

De las clasificaciones encontradas, la de Lykken (Lykken, 1968) es la más antigua. Esta clasificación aparece publicada en el año 1968, mientras que las más recientes son las clasificaciones de Schmidt (Schmidt, 2009) así como la de Evanschitzky y Armstrong (Evanschitzky y Armstrong, 2010) publicadas en 2009¹.

La clasificación más citada es la de Lykken (Lykken, 1968), seguida de la clasificación de Hendrick (Hendrick, 1990). Las clasificaciones de Lykken (Lykken, 1968), Hendrick (Hendrick, 1990) y Bahr et al. (Bahr et al., 1983) han sido citadas no sólo en sus propias disciplinas sino también en algunas áreas de negocios. Las citas que se han tenido en cuenta son aquellas donde el autor hace referencia en alguna manera a la clasificación, por lo que se omitieron citas a otras cuestiones de los artículos donde se describen las clasificaciones. En la tabla 2.3 se muestran las clasificaciones agrupadas por las veces que han sido citadas, sólo se presentan las clasificaciones citadas una o más veces.

¹Aparecida en la versión electrónica del respectivo journal en 2009 e impresa en el año 2010.

Área	Número	Autor(es)
-Ciencias Sociales-		
Psicología	5	Lykken; Hendrick; Hunter; Schmidt; Kantowitz et al.
Sociología	3	Finifter; La Sorte; Bahr et al.
Economía	1	Mittelstaedt & Zorn
Comunicación Humana	1	Kelly et al.
Desarrollo Humano	1	Van IJzendoorn
-Negocios-		
Mercadotecnia	3	Leone & Schultz; Easley et al.; Monroe
Contabilidad	1	Lindsay & Ehrenberg
Gestión	1	Tsang & Kwan
Predicción	1	Evanschitzky & Armstrong
-Filosofía-		
Filosofía de la Ciencia	1	Radder
TOTAL	18	

Tabla 2.2: Clasificaciones de replications por área

Clasificación	Número de veces citada	Citada por
Lykken	7	Finifter; Kelly et al.; Bahr et al; Mittelstaedt & Zorn; Van IJzendoorn; Easley et al.; Schmidt
Hendrick	3	Tsang & Kwan; Easley et al.; Schmidt
La Sorte	2	Bahr et al.; Van IJzendoorn
Bahr et al.	2	Van IJzendoorn; Tsang & Kwan
Lindsay & Ehrenberg	2	Tsang & Kwan; Schmidt
Finifter	1	Bahr et al.
Mittelstaedt & Zorn	1	Tsang & Kwan
Leone & Schultz	1	Easley et al.
Monroe	1	Easley et al.
Radder	1	Schmidt

Tabla 2.3: Clasificaciones más citadas

Las búsquedas antes descritas se utilizaron en Gómez et al. (Gómez et al., 2010b). Una de las observaciones hechas en este trabajo fue referente a la forma de hacer las búsquedas. En éstas sólo se usó el término *replicación* por lo que se sugirió también usar el término *reproducción* con el fin de encontrar más tipos de replications. Tomando en cuenta las observaciones se realizó de nuevo la búsqueda de clasificaciones de manera sistemática.

2.2.2 Segunda Búsqueda de Replicaciones

Como punto de partida se decidió utilizar la base de datos Scopus[®] para realizar las búsquedas de las clasificaciones. Se seleccionó esta base de datos porque tiene un amplio nivel de cobertura de publicaciones en diferentes disciplinas científicas y porque el motor de búsqueda que utiliza es amigable y, ante todo, robusto (Dieste et al., 2009).

Una vez elegida la base de datos, se eligieron frases en inglés que puedan emplearse para distinguir diferentes formas de realizar una replicación como son: *classification of replications*, *kinds of replication*, *replication kinds*, *types of replication*, *replication types*, *typology of replications*, *replications typology*, *taxonomy of replications* así como *replica-*

tions taxonomy.

Tomando como referencia el conjunto de frases identificadas, se eligieron los términos: *replication*, *classification*, *kind*, *type*, *typology* y *taxonomy* como términos clave para usar en la cadena de búsqueda.

La cadena se definió de tal forma que Scopus[®] busque el término *replication* junto con alguno de los términos restantes en títulos de artículos, sumarios y palabras clave de la base de datos i.e. TITLE-ABS-KEY(replication AND (classification OR kind OR type OR typology OR taxonomy)). Los términos se especificaron en singular ya que Scopus[®] los busca automáticamente tanto en singular como en plural. Como resultado de esta cadena de búsqueda se obtuvieron 46,783 documentos.

Dado el gran número de documentos que arrojó la búsqueda, se decidió refinar la búsqueda dividiéndola en cinco cadenas donde se usaron sólo dos términos, el término *replication* más uno de los cinco términos restantes (*classification*, *kind*, *type*, *typology* y *taxonomy*). En la tabla 2.4 se muestra el número de documentos resultantes de estas nuevas cadenas de búsqueda.

Búsqueda	Cadena de Búsqueda	Documentos resultantes
1	TITLE-ABS-KEY(replication AND classification)	2,541
2	TITLE-ABS-KEY(replication AND kind)	1,290
3	TITLE-ABS-KEY(replication AND type)	43,592
4	TITLE-ABS-KEY(replication AND typology)	80
5	TITLE-ABS-KEY(replication AND taxonomy)	289

Tabla 2.4: Número de documentos resultantes

Con excepción de la búsqueda número 3, usar cadenas más cortas (en este caso de dos términos) facilita considerablemente la identificación de los documentos relevantes pues permite una agrupación preliminar de los resultados de la búsqueda por relevancia dependiendo de la especificidad de los términos. La búsqueda número 3 arroja un mayor número de resultados debido a que el término *type* es más común (y por tanto menos adecuado como palabra clave) que el resto de par de términos.

De acuerdo como se especificaron las cadenas de búsqueda, Scopus[®] localiza el par de términos en títulos, sumarios y palabras clave de los documentos sin importar la distancia en que se encuentren los términos, esto implica que muchos de los documentos resultantes no mencionen o hablen sobre las diferentes formas de realizar una replicación. Aunque se es consciente de esta situación, parece merecer la pena examinar los títulos, sumarios y palabras clave de las búsquedas 1, 2, 4 y 5, dado su tamaño razonable; No así, los resultados de la búsqueda 3.

Para tratar los resultados de la cadena 3 se decidió refinarla restringiendo la distancia entre los términos *replication* y *type* de modo que pudieran encontrarse documentos más cercanos al objetivo deseado. Para esto se usó el operador de proximidad *Within* que busca un par de términos dentro de una distancia determinada sin importar el orden en que se encuentren ubicados i.e. cualquiera de los términos puede aparecer en primer lugar. En esta nueva cadena se definió una distancia de dos términos entre la palabras *replication* y *type*, por lo que la cadena resultante se especificó como TITLE-ABS-KEY(replication W/2 type). Con esta nueva cadena se obtuvieron 3,143 documentos, un número más tratable para realizar la revisión de títulos, sumarios y palabras clave en busca de artículos relevantes.

Conforme se realizó el proceso de revisión de títulos, sumarios y palabras clave de los 7,343 documentos resultantes de las cinco cadenas usadas, se separaron los documentos que podrían discutir diferentes formas de realizar una replicación.

A continuación se localizó y examinó el contenido de los documentos relevantes identificados durante la revisión de los resultados de la búsqueda. Tras leer los documentos relevantes, se identificaron 9 documentos que describen tipos de replicaciones. En la mayoría de estos documentos, los autores citan otros trabajos que describen diferentes tipos de replicaciones. De manera conjunta se identificaron 25 clasificaciones de replicaciones, 9 a través de Scopus[®] y 16 a través de las referencias mencionadas en los documentos. De estas 25 clasificaciones de replicaciones se reunieron 20 para su análisis. En el apéndice A se encuentran las descripciones de los diferentes tipos de replicaciones encontrados.

En la tabla 2.5 se muestran las clasificaciones identificadas con las cadenas de búsqueda utilizadas y se indica si lograron obtenerse ó no.

Cadenas de búsqueda usando el código TITLE-ABS-KEY	Clasificación	Clasificaciones citadas	Documento Obtenido
replication AND classification	Beck	Finifter (1975)	✓
		Lykken	✓
		Sidman	✓
		La Sorte	✓
		Blomquist	✓
replication AND kind	Lüdtke	-	-
replication W/2 type	Schmidt	Radder	✓
		Lykken	✓
		Hendrick	✓
		Sargent	x
		Keppel	x
		Beloff	x
		Rao	x
		Adams et al.	✓
	Tsang and Kwan	Kelly et al.	✓
		La Sorte	-
		Brown and Coney	✓
		Fuess	✓
		Hendrick	✓
		Lindsay and Ehrenberg	✓
		Mittelstaedt and Zorn	✓
		Reid et al.	x
replication AND typology	Bahr et al.	Lykken	✓
		Finifter	✓
		La Sorte	✓
		Hyman and Wright	✓
replication AND taxonomy	Kelly et al.	Sidman	✓
		Lykken	✓
		Barker and Gurman	✓
		Finifter	✓
	Finifter (1975)	Finifter	✓

Tabla 2.5: Citas de las clasificaciones identificadas

Respecto a las 5 clasificaciones que no se incluyeron para su análisis, la clasificación de Reid et al. (Reid et al., 1981) no fue posible localizarla en la página del journal ni en el repositorio de archivos históricos que menciona el journal. Respecto a las clasificaciones de Sargent (Sargent, 1981), Keppel (Keppel, 1982), Beloff (Beloff, 1985) y Rao (Rao, 1981) se obtuvo la primera clasificación (Sargent, 1981) y después de leer el artículo se decidió no localizar las otras tres y descartar esta clasificación debido a que estas clasificaciones pertenecen al área de parapsicología, área conflictiva para ser incorporada

una investigación científica.

Aunque *replication* es el término de base usado en estas búsquedas, los términos *reproduction* y *repetition* pueden usarse para denotar una replicación. Con el fin de localizar más clasificaciones de replicaciones que utilicen estos términos, se realizaron de nuevo las búsquedas con las mismas cadenas empero sustituyendo el término *replication* por *reproduction* y *repetition* respectivamente. En la tabla 2.6 se muestra el número de documentos resultantes de las cadenas de búsqueda que incluyen los términos *reproduction* y *repetition*.

Búsqueda	Cadenas de Búsqueda	Documentos Resultantes
6	TITLE-ABS-KEY(reproduction AND classification)	2,381
7	TITLE-ABS-KEY(reproduction AND kind)	1,888
8	TITLE-ABS-KEY(reproduction AND type)	14,759
9	TITLE-ABS-KEY(reproduction W/2 type)	466
10	TITLE-ABS-KEY(reproduction AND typology)	141
11	TITLE-ABS-KEY(reproduction AND taxonomy)	819
12	TITLE-ABS-KEY(repetition AND classification)	694
13	TITLE-ABS-KEY(repetition AND kind)	1,034
14	TITLE-ABS-KEY(repetition AND type)	5,902
15	TITLE-ABS-KEY(repetition W/2 type)	266
16	TITLE-ABS-KEY(repetition AND typology)	88
17	TITLE-ABS-KEY(repetition AND taxonomy)	70

Tabla 2.6: Cadenas de búsqueda utilizadas con el término *reproduction*

De igual modo que se refinó la cadena *replication AND type* con el operador de proximidad *Within*, se refinaron las cadenas *reproduction AND type* y *repetition AND type*. Tras examinar los títulos, sumarios y palabras clave de los documentos resultantes, no se encontraron trabajos que mencionen diferentes formas de realizar una replicación, por ejemplo, con los diferentes términos (*classification*, *kind*, *type*, *typology* y *taxonomy*) se observa que el término *reproduction* a menudo se usa para referirse al proceso por el que los organismos vivos dan lugar a la aparición de nuevos organismos, en menor medida, el término *repetition* se utiliza en áreas afines a la lingüística. Ambos términos se utilizan en mayor o menor medida en las diferentes disciplinas científicas.

Ciertos autores emplean los términos *replication*, *reproduction* and *repetition* de manera indistinta. En esta tesis se usa cada uno de estos términos como diferentes formas de replicar un experimento. Se usa “replicar” para referirnos a la repetición de un experimento con el fin de comprobar sus resultados y no “reproducir” o “repetir”. Replicar es el término más extendido en la mayoría de las ciencias para referirse a la repetición de un experimento tal y como se ha contrastado empíricamente realizando búsquedas en la base de datos Scopus® con los tres términos.

Respecto a las clasificaciones obtenidas en esta segunda búsqueda, la clasificación más citada es la de Lykken (Lykken, 1968) (citada cuatro veces), seguida de las clasificaciones de Finifter (Finifter, 1972) (citada tres veces), Hendrick (Hendrick, 1990) (citada dos veces), La Sorte (La Sorte, 1972) (citada dos veces) y Sidman (Sidman, 1960) (citada dos veces). Esta última clasificación es la más antigua de las encontradas tanto en la primera como en la segunda búsqueda.

2.3 Resultados de ambas Búsquedas

En total se obtuvieron 27 clasificaciones de replications, habiendo en común entre los resultados de las dos búsquedas 11 clasificaciones. Del total de clasificaciones, 15 (56%) pertenecen a áreas de ciencias sociales, 8 (30%) al área de negocios, 3 (11%) al área de salud y 1 (4%) al área de filosofía. En la tabla 2.7 se listan los autores de las clasificaciones, el área, el año de aparición de la clasificación y si ésta se encontró en la primera, segunda o en ambas búsquedas.

En la mayoría de las clasificaciones los autores buscan originalidad al nombrar los tipos de replications y por tanto usan sus propios términos para denominar un tipo de replicación. Existen algunas excepciones como Kelly et al. (Kelly et al., 1979) o Adams et al. (Adams et al., 2005) que emplean los mismos términos de la clasificación de Lykken (Lykken, 1968). En las clasificaciones encontradas también se observa que no hay una estandarización intra-disciplina o inter-disciplinas para nombrar los diferentes tipos de replicación.

En algunos casos más de un autor coincide con el mismo término. Pero aunque los autores coincidan en los términos, esto no implica que denoten el mismo tipo de replicación. El significado de un término entre autores puede ser igual, similar o distinto. Por ejemplo, Easley et al. (Easley et al., 2000) así como Mittelstaedt y Zorn (Mittelstaedt y Zorn, 1984) coinciden en algunos términos, pero los tipos de replicación a los que se refieren son diferentes. Otro ejemplo, las definiciones de *replicación conceptual* descritas por Hunter (Hunter, 2001) y Monroe (Monroe, 1992) son similares entre sí. También son equivalentes entre sí las definiciones de *replicación conceptual* descritas por Hendrick (Hendrick, 1990), Kantowitz et al. (Kantowitz et al., 1984) y Schmidt (Schmidt, 2009).

Autor(es)	Área	Año	Búsqueda
Adams et al.	Sociología	2005	2da
Bahr et al.	Sociología	1983	1ra y 2da
Barker and Gurman	Psicología	1972	2da
Beck	Enfermería	1994	2da
Blomquist	Enfermería	1986	2da
Brown and Coney	Mercadotécnica	1976	2da
Easley et al.	Mercadotécnica	2000	1ra
Evanschitzky and Armstrong	Predicción	2009 y 2010	1ra
Finifter	Sociología	1972 y 1975	1ra y 2da
Fuess	Negocios	1996	2da
Hendrick	Psicología	1990	1ra y 2da
Hunter	Psicología	2001	1ra
Hyman and Wright	Sociología	1967	2da
Kantowitz et al.	Psicología	1984	1ra
Kelly et al.	Comunicación Humana	1979	1ra y 2da
La Sorte	Sociología	1972	1ra y 2da
Leone and Schultz	Mercadotécnica	1980	1ra
Lindsay and Ehrenberg	Contabilidad	1993	1ra y 2da
Lüdtke	Medicina	2008	2da
Lykken	Psicología	1968	1ra y 2da
Mittelstaedt and Zorn	Economía	1984	1ra y 2da
Monroe	Mercadotécnica	1992	1ra
Radder	Filosofía de la Ciencia	1992	1ra y 2da
Schmidt	Psicología	2009	1ra y 2da
Sidman	Psicología	1960	2da
Tsang and Kwan	Gestión	1999	1ra y 2da
Van IJzendoorn	Desarrollo Humano	1994	1ra

Tabla 2.7: Clasificaciones encontradas en las dos búsquedas

Sin embargo, estas definiciones de *replicación conceptual* son diferentes respecto a las definiciones de Hunter (Hunter, 2001) y Monroe (Monroe, 1992). De acuerdo a Hunter (Hunter, 2001) y Monroe (Monroe, 1992) en la replicación conceptual se varían los procedimientos, las variables independientes y se examinan otras condiciones que puedan influir en los resultados. Sin embargo, de acuerdo a Hendrick (Hendrick, 1990), Kantowitz et al. (Kantowitz et al., 1984) y Schmidt (Schmidt, 2009) en una replicación conceptual se intenta reproducir un resultado de una manera completamente distinta a la del experimento original. Siguiendo con las definiciones de *replicación conceptual* descritas por Hendrick (Hendrick, 1990), Kantowitz et al. (Kantowitz et al., 1984) y Schmidt (Schmidt, 2009), otros autores usan términos equivalentes a esta replicación, por ejemplo, Lykken (Lykken, 1968) la llama como *replicación de constructos* (*constructive replication*) mientras que La Sorte (La Sorte, 1972) la define como *replicación teórica*.

En base a los tipos de replications encontrados, se vislumbran tres grupos de replications:

1. Replicaciones que no varían o varían ligeramente respecto al experimento base
2. Replicaciones que varían empero siguen el mismo protocolo experimental usado en el experimento base
3. Replicaciones que utilizan diferentes protocolos experimentales para verificar los resultados del experimento base.

En la tabla 2.8 se muestran los distintos tipos de replications organizados en estos tres grupos.

En los primeros dos grupos, el nivel de similitud entre la replicación y el experimento base puede variar. Esta variación implica que en la replicación se cambie o se modifique en cierta medida alguno de los elementos de la estructura del experimento. En caso del primer grupo la replicación se mantiene lo más similar posible al experimento base mientras que en el segundo grupo la replicación varía respecto al experimento base. En ambos grupos las replications siguen el mismo protocolo experimental usado en el experimento base.

En el tercer grupo, los investigadores emplean un protocolo experimental distinto al del experimento base para verificar los resultados. En esta replicación lo único que hay en común con el experimento base es que ambos experimentos comparten una misma estructura teórica, es decir, comparten los mismos constructos e hipótesis a estudiar. Esta replicación se utiliza para verificar los resultados previamente observados a través de un protocolo experimental distinto.

Se encontraron otras dos formas de repetir un experimento que pueden o no considerarse como una replicación, la primera forma se presenta cuando los investigadores utilizan observaciones existentes para reanalizarlas de nuevo, ya sea utilizando procedimientos iguales o diferentes a los del experimento base. Ejemplos de esta forma de repetición son: *checking of analysis, reanalysis of data* (Tsang y Kwan, 1999), *internal replication* (La Sorte, 1972), *pseudoreplication* (Finifter, 1972) así como *types I and II* (Mittelstaedt y Zorn, 1984). La segunda forma de repetición se presenta cuando se varían y comparan los modelos estadísticos respecto a los modelos usados en el estudio base. Ejemplos de estas formas de repetición son: *types II y IV* (Mittelstaedt y Zorn, 1984) y *Model Comparisons* (Evanschitzky y Armstrong, 2010). Estas dos formas de repetición se consideran como re-análisis de datos (Gómez et al., 2010a) y no una verdadera replicación, puesto que no existe una repetición del protocolo experimental. En este sentido la clasificación de Mittelstaedt y Zorn (Mittelstaedt y Zorn, 1984) no aparece en ninguno de los tres grupos de replications porque se asemeja más a una forma de re-análisis de

Autor(es)	Grupo I	Grupo II	Grupo III
Adams et al.	Literal	Operational, Instrumental	Constructive
Bahr et al.	Types A..D	Types E..H	Types I..P
Barker and Gurman	Type I	Types II and III	Type IV
Beck	Type I	Type II	Type III,V
Blomquist	Literal		Construct
Brown and Coney		Replication and Replication with extension	
Easley et al.	Types 0 and I	Type II	Type III
Evanschitzky and Armstrong	Real		
Finifter		Virtual	Systematic
Fuess	Duplication	Replication with extension	
Hendrick	Strict	Partial	Conceptual
Hunter	Statistical	Conceptual, Scientific	
Hyman and Wright	Type II	Type I	Type III
Kantowitz et al.	Direct	Systematic	Conceptual
Kelly et al.	Literal	Operational, Instrumental	Constructive
La Sorte		Retest, Independent	Theoretical
Leone and Schultz	Experimental	Nonexperimental	Corroboration
Lindsay and Ehrenberg	Close	Differentiated	Differentiated
Lüdtke	Exact	Independent	Independent
Lykken	Literal	Operational	Constructive
Monroe		Conceptual, Demonstrated, Partial Independent, Nonindependent Sequential	
Radder	Reproducibility of the material realization of an experiment	Reproducibility of an experiment under a fixed theoretical interpretation	Reproducibility of the result of an experiment
Schmidt	Direct	Direct	Conceptual
Sidman	Direct	Systematic	
Tsang and Kwan	Exact	Empirical Generalization, Generalization and Extension	Conceptual
Van IJzendoorn	Exact	Varied	

Tabla 2.8: Grupos de replicaciones identificados

datos.

De acuerdo a los grupos de replications identificados, se observa que una replicación puede tener diferentes niveles de similitud respecto al experimento base. Es decir, en la replicación se pueden variar distintos elementos de la estructura del experimento. En la tabla 2.9 se muestran algunos elementos experimentales que las clasificaciones estudiadas permiten variar en una replicación. Nótese que dependiendo de la variación el tipo de replicación varía.

Elemento variable	Tipo de replicación	Autor(es)
Instrumentos de medición	Differentiated Replication	Lindsay & Ehrenberg
Medidas	Operational Replication	Kelly et al.
Método	Conceptual Replication	Schmidt
Lugar	Types B,F,J,N,D,H,L,P	Bahr et al.
Poblaciones	Empirical Generalization	Tsang & Kwan
Diseño de la investigación	Retest Replication	La Sorte
Investigador	Independent Replication	Monroe
Muestra	Virtual Replication	Finifter

Tabla 2.9: Elementos identificados que pueden variar en la replicación

Aunque el objetivo general de la replicación es la verificación de resultados, se observa que los diversos tipos de replicación tienen diferentes objetivos o propósitos específicos. Por ejemplo, de acuerdo a Lykken (Lykken, 1968) el propósito de la *operational replication* es que la receta experimental usada produzca los mismos resultados con otro investigador. Sin embargo, la *systematic replication* de Finifter (Finifter, 1972) tiene como propósito producir nuevos hallazgos usando procedimientos distintos a los del experimento base.

Capítulo 3

Planteamiento del Problema

RESUMEN: En este capítulo se profundiza en la descripción del problema específico que esta investigación aborda, así como el tipo de aproximación seguido en la elaboración de la solución.

3.1 Descripción del Problema

Un componente fundamental del paradigma experimental es la replicación. Para construir y consolidar un cuerpo de conocimientos a partir de resultados experimentales es necesario verificar ampliamente dichos resultados. Esta verificación se realiza mediante repeticiones de un experimento. Los experimentos arrojan resultados que se deben poder reproducir. Replicando un experimento es posible comprobar si se reproducen, o no, los eventos previamente observados. Mediante la reproducción de un mismo evento en diversas repeticiones se identifican regularidades existentes en la realidad bajo estudio (en nuestro caso la construcción de software). Conocer tales regularidades permite identificar los mecanismos que regulan el evento bajo estudio o, como mínimo, predecir su comportamiento.

Hoy en día en ISE la mayoría de los resultados experimentales son aislados, es decir, la mayoría de resultados experimentales no se han corroborado. Sjøberg et al. (Sjøberg et al., 2005) realizaron un estudio para conocer un estimado del número de experimentos controlados reportados en la literatura. Los autores extrajeron una muestra de 5,453 artículos publicados en diferentes journals y actas de conferencias afines a IS. Los autores encontraron un total de 113 experimentos controlados, de los cuales sólo 20 (17.7%) se describen como repeticiones.

Por otra parte, Zannier et al. (Zannier et al., 2006) realizaron un estudio para conocer la cantidad de artículos que contienen estudios empíricos publicados durante 29 años que se han realizado las conferencias del ICSE (*International Conference on Software Engineering*). De una población de 1,227 artículos, los autores extrajeron una muestra de 63 (5%) artículos, de esa muestra los autores no encontraron ningún estudio que pudiera considerarse replicación.

Respecto al Journal en Ingeniería del Software Empírica, se ha realizado una búsqueda para conocer el número de estudios replicados. Desde la aparición de este journal en 1996 hasta la fecha¹ se encontraron únicamente nueve replications².

Si un experimento no es replicado no hay forma de distinguir entre las siguientes situaciones:

- Los resultados se han producido de manera fortuita (el evento se produjo por casualidad)
- Los resultados son artificiales (el evento ocurre únicamente en el experimento, no en la realidad)
- Los resultados corresponden verdaderamente a una regularidad de la realidad

Una replicación necesita ciertos elementos comunes respecto a su experimento base. Cuando un fenómeno o evento comienza a estudiarse, experimentalmente existen muchos aspectos que se desconocen. Introducir algún cambio en la replicación por pequeño que éste sea puede ocasionar una variación en los resultados del experimento imposible de explicar. En las disciplinas experimentales inmaduras comenzar con replications lo más similares posibles permite conocer las condiciones experimentales que deben controlarse (Brinberg y McGrath, 1985). Cuando un evento ya es bien conocido existe un alto control sobre las condiciones experimentales que influyen en los resultados de un experimento. Mediante replications menos similares se descartan los resultados artificiales. Es decir, verificar los resultados observados con distintos protocolos experimentales permite identificar eventos independientes del experimento.

En ISE, la realización de replications ha sido complicada debido al estado de inmadurez de la disciplina. Dado el desconocimiento de los mecanismos que regulan la construcción de software así como las condiciones experimentales relevantes para estudiarlo, cualquier cambio en la replicación por mínimo que sea puede ocasionar alteraciones inexplicables en los resultados. Sin embargo, a menudo los investigadores se ven obligados a realizar modificaciones en el contexto experimental. Estas modificaciones pueden ocasionar variaciones importantes en los resultados de la replicación que hacen difícil verificar los resultados del experimento base. En varios intentos de combinar resultados de replications ISE Hayes (Hayes, 1999), Miller (Miller, 1999, 2000, 2005), Hannay et al. (Hannay et al., 2009), Jørgensen (Jørgensen, 2004), Pickard et al. (Pickard et al., 1998), Shull et al. (Shull et al., 2002) y Juristo et al. (Juristo et al., 2004) han reportado que los resultados de las replications son tan diferentes que resulta complicado agregarlos. Se han introducido tantos cambios que resultaba imposible extraer consecuencias de la comparación de los resultados.

La inmadurez experimental en ISE lleva a una incomprensión del concepto de replicación aplicado a los experimentos de IS. Cuando se piensa en la replicación se tiene como estereotipo de replicación aquella que se realiza independientemente por otros investigadores en sitios distintos a los del experimento base. Sin embargo este estereotipo no se corresponde con algunas de las replications en ISE realizadas o de manera conjunta o reutilizando parte de los materiales empleados en el experimento base o en el mismo sitio (Basili y Selby, 1987; Porter et al., 1995; Fusaro et al., 1997; Miller et al., 1998; Sandahl et al., 1998; Porter y Votta, 1998; Wood et al., 1997; Juristo y Vegas, 2003; Vegas et al.,

¹Búsqueda realizada el 18/May/2011.

²El criterio de búsqueda se realizó en la URL del journal <http://www.springerlink.com/content/1382-3256>. La búsqueda se realizó sólo en los títulos y sumarios de los artículos. El criterio de búsqueda empleado fue el siguiente: ti:(replication) OR su:(replication).

2006; Juristo y Vegas, 2009). Esta situación pareciera ocasionar confusión sobre si estas replications violan o no el principio de independencia.

Existe cierta controversia en ISE respecto a la forma de realizar replications, hay quienes proponen replicar experimentos reutilizando materiales del experimento base (Basili et al., 1999; Shull et al., 2008) para replicar lo más similar posible, por el contrario, hay quienes proponen usar protocolos y materiales distintos a los del experimento base (Kitchenham, 2008; Miller, 2005) para preservar el principio de independencia y evitar propagar errores en replications que usen los mismos materiales. Otros sugieren formas alternativas de verificación de resultados (Miller, 2008) para entender la problemática de la replicación en ISE. Estas controversias parecen deberse a que en ISE aún no se ha adaptado adecuadamente el principio de la replicación del paradigma experimental.

Aunque el fin último de la replicación sea verificar un resultado, las replications pueden tener propósitos específicos de verificación (por ejemplo, verificar si los resultados pueden reproducirse de nuevo en el mismo sitio con los mismos experimentadores, en otros sitios, con otros materiales, o con diferentes tipos de poblaciones). A efectos de verificación de resultados, una replicación debe ser exacta; pero a efectos de búsqueda de conocimiento y aprendizaje una replicación debe tener algunas diferencias. Ambas son necesarias. En una fase temprana de investigación experimental, como ocurre en ISE, la igualdad ni siquiera es posible; más adelante, tanto la búsqueda como la prueba pueden ser más sistemáticas. En las etapas precoces la no consecución de los resultados esperados en una replicación no debe entenderse como fracaso, sino como un paso hacia el descubrimiento de algún factor nuevo (Brinberg y McGrath, 1985). El papel de la replicación es todavía mal entendido en ISE. Se necesita más investigación sobre la replicación como generadora de aprendizaje de nuevas variables relevantes involucradas en la construcción de sistemas software.

En ISE se han hecho algunos intentos de explotación de replications como son la agregación o combinación de resultados experimentales a través de técnicas estadísticas (Hayes, 1999; Miller, 1999, 2000, 2005; Hannay et al., 2009; Jørgensen, 2004; Pickard et al., 1998; Shull et al., 2002), los autores han reportado que los resultados de las replications son tan diferentes que resulta complicado agregarlos. La aplicación de técnicas estadísticas de agregación es sólo el primer paso del trabajo que debe (y puede) hacerse al comparar resultados experimentales. En disciplinas experimentalmente inmaduras como la ISE la agregación de replications puede ayudar a comprender mejor el contexto donde se realizan los experimentos mediante el aprendizaje de nuevas variables relevantes.

3.2 Hipótesis de Trabajo

De manera general el problema de investigación que se aborda en esta tesis es la comprensión y adaptación del concepto de replicación a la ISE. Para abordar este problema se definen las siguientes hipótesis de trabajo.

H_1 El concepto de replicación del paradigma experimental es adaptable a la ISE. Estudiar el concepto de replicación ayudaría a comprender mejor su utilización en ISE.

H_2 Es viable contar con una tipología de replications que oriente al investigador sobre qué tipo de replicación se ajusta a sus condiciones. Contar con una tipología de replications para la ISE ayudaría a comprender mejor el objetivo particular de cada tipo de replicación así como los cambios que implican cada uno.

H_3 Es posible mejorar la explotación de conjuntos de replicaciones en ISE para conocer el grado de verificación logrado de los resultados experimentales. Aunque de manera no satisfactoria se han hecho algunos intentos de explotación de conjuntos de replicaciones como son la agregación o combinación de resultados experimentales, no se han desarrollado ideas clave que permitan al investigador comparar de manera sistemática replicaciones, entender el grado de verificación logrado en los resultados y empaquetar este conocimiento para que sea útil tanto para el profesional como para el investigador.

3.3 Estructura de la Solución Propuesta

Para abordar el problema planteado se necesita realizar tres tareas interrelacionadas pero bien diferenciadas. La solución de la problemática a tratar se compone de tres partes que son:

1. Estudio del concepto de replicación
2. Desarrollo de una tipología de replicaciones para la ISE
3. Desarrollo de un marco conceptual con ideas clave para realizar comparaciones entre replicaciones

Para clarificar la incomprensión y malentendido del concepto de replicación en ISE, se hace necesario realizar un trabajo de investigación sobre el significado y las implicaciones del concepto de replicación en otras disciplinas científicas.

A raíz de la investigación de la primera tarea, se desarrollará una clasificación de replicaciones que clarifique los tipos de replicación que pueden llevarse a cabo en ISE, los cambios permitidos en cada tipo de replicación así como el rol de verificación que acarrea cada tipo. La tipología permitirá orientar al investigador a identificar qué tipo de replicación es susceptible de realizar y qué grado de verificación podrá alcanzar con ella.

La tercera tarea de la solución se desarrolla un marco conceptual para la comparación de replicaciones. Este marco facilitará la comparación de conjuntos de replicaciones para obtener conocimiento de ellas. Este marco se conformará por varias ideas clave que ayudarán a:

- Identificar las unidades comparables comunes entre varias replicaciones.
- Comparar replicaciones no sólo a nivel de resultados experimentales sino también a nivel de cambios de los distintos elementos que constituyen las replicaciones.
- Cuando los resultados de las replicaciones sean reproducibles, este marco permitirá conocer el grado de verificación alcanzado dependiendo de los cambios identificados en las replicaciones.
- Explorar variables moderadoras que ocasionan resultados no reproducibles. La exploración de estas variables contribuirá al aprendizaje de las distintas variables que intervienen en la construcción de software.

Finalmente, siguiendo las ideas de este marco conceptual, como cuarta tarea se desarrollará un procedimiento que facilite la comparación de conjuntos de replicaciones.

Parte II

Solución

Capítulo 4

Clarificación de Conceptos de Replicación para la ISE

RESUMEN: En este capítulo se realiza un estudio del concepto de replicación con el fin de aliviar su incomprensión en ISE.

Como primer paso hacia la solución es necesario clarificar algunos conceptos de replicación mal comprendidos por ISE. Aunque esta clarificación no es contribución totalmente original de esta tesis porque se basa en conocimiento existente en otras disciplinas, se ha realizado un trabajo de síntesis referente al uso de la replicación en otras ciencias con el fin de importar y clarificar conceptos de utilidad para la ISE. El trabajo de síntesis realizado ha resultado especialmente complejo porque no existen libros específicos que aborden el tema de la replicación en ciencia y que resuelvan para otras disciplinas los problemas planteados en esta investigación.

4.1 Repetición, Replicación y Reproducción

En las distintas clasificaciones de replications encontradas, la mayoría de los autores usa el término *replicación* para referirse a la repetición de un experimento, sin embargo algunos autores usan el término *reproducción* o *reproducibilidad* para referirse a tal repetición. Algunos autores utilizan ambos términos como sinónimos. También Wikipedia utiliza estos términos indistintamente al definir la *reproducibilidad* como, “uno de los principios esenciales del método científico, y se refiere a la capacidad que tenga una prueba o experimento de ser reproducido o replicado” (Wikipedia, 2009).

Sin embargo, algunos autores hacen una distinción entre ambos términos. Por ejemplo, Cartwright (Cartwright, 1991) propone distinguir *replicabilidad* “realizar el mismo experimento otra vez” de *reproducibilidad* “realizar un nuevo experimento”. Para Cartwright (Cartwright, 1991) la replicación de un experimento se refiere a repetir un experimento siguiendo muy de cerca el protocolo utilizado en el experimento previo, mientras que la reproducción se refiere a repetir la observación de un evento previamente observado utilizando un protocolo diferente al del experimento previo.

De acuerdo a Cartwright (Cartwright, 1991) la replicación no garantiza que el resultado observado represente la realidad bajo estudio. El resultado puede ser artificial; es decir, producto de los materiales o los instrumentos empleados en el experimento. Si se desea garantizar que un resultado experimental corresponde a la realidad bajo estudio, entonces se debe realizar una reproducción, empleando diferentes protocolos experimentales para asegurarse que el evento observado es independiente del protocolo, materiales o instrumentos usados en el experimento base.

Cuando los resultados son repetibles con el mismo protocolo experimental, los investigadores tienen la confianza de que han observado algún tipo de evento lo suficientemente estable como para ser observado múltiples veces, lo suficientemente objetivo para ser observado por diversos investigadores y lo suficientemente global para ser observado en diferentes laboratorios. Pero al haber sido siempre observado con el mismo protocolo, la relación protocolo-evento puede ser muy estrecha. En palabras de Radder (Radder, 1996), “Este resultado no implica algún acuerdo respecto al evento que es. Incluso es posible que algunos interpretes discutan que el evento es artificial, porque, aunque éste sea estable, no es atribuido al objeto bajo estudio sino a ciertas características del aparato¹”.

En esta situación, Cartwright (Cartwright, 1991) afirma que la *reproducibilidad* “es entonces una protección en contra de fallos en nuestros instrumentos”. Según Cartwright (Cartwright, 1991) cuanta mayor seguridad se tenga en el diseño del protocolo, menor es la necesidad de emplear la *reproducibilidad*.

La *reproducción* puede ser vista como una forma de triangulación, en la que los investigadores emplean distintos protocolos en un intento de validar o corroborar los resultados encontrados en el experimento base (Easterbrook et al., 2008). De acuerdo a Park (Park, 2004), “estas estrategias de triangulación pueden ser usadas para sustentar un hallazgo conceptual, pero no son replicaciones en ningún grado”.

El término *reproducción* es también usado en teoría de la medida con un significado análogo. En ISO 5725-1 (ISO, 1994) se definen las *condiciones de reproducibilidad* de una medida como condiciones en que los resultados de un ensayo se obtienen aplicando el mismo método a elementos de ensayo idénticos, en laboratorios distintos, con distintos operadores y utilizando equipos diferentes. En IS Anda et al. (Anda et al., 2009) utilizan la connotación de este término en un caso de estudio múltiple para evaluar el nivel de reproducibilidad de varios proyectos software.

Un uso similar de este término se encuentra en ciencias de la computación donde existe una iniciativa denominada *reproducible research* (Kovacevic, 2007; Fomel y Claerbout, 2009) la cual se refiere a la idea de que cualquier cosa en un artículo científico debe ser reproducible por el lector, esto es, resultados, códigos o gráficos. En nuestra disciplina, Meyer (Meyer, 2009) ha empleado el término *reproducibility* con significado análogo al de la iniciativa *Reproducible Research*.

Repetición es otro término relacionado con la replicación. En teoría de la medida (ISO, 1994) se definen las *condiciones de repetibilidad* de una medida como condiciones en que los resultados de un ensayo se obtienen mediante la aplicación del mismo método por el mismo operador a elementos de ensayo idénticos y en el mismo laboratorio, utilizando el mismo equipo a breves intervalos de tiempo. Wikipedia (Wikipedia, 2009) menciona que la *repetibilidad* “mide la tasa de éxito en experimentos sucesivos, posiblemente conducidos por los mismos experimentadores”. En este sentido la repetición es utilizada por los investigadores para aumentar el tamaño de la muestra y así garantizar que el efecto

¹El término aparato se refiere a los instrumentos, materiales o métodos empleados y que en conjunto conforman el protocolo experimental.

observado no es producto del error muestral.

La repetición puede ser vista como una forma de replicación ya que ambos mecanismos siguen un mismo protocolo para verificar un resultado. La diferencia radica en que en la repetición un mismo investigador repite sus propios experimentos, mientras que en la replicación un investigador repite el experimento de otros. La repetición logra alcanzar un nivel de similitud mucho mayor ya que el experimento se repite en el mismo sitio por el mismo investigador con el mismo protocolo. En una replicación el nivel de similitud varía dependiendo del número de cambios que el nuevo investigador realiza en el protocolo (deliberada o indeliberadamente).

Como se observa la repetición, replicación y reproducción conllevan propósitos particulares. En base a esto se propone identificar en ISE a la repetición, replicación y reproducción como formas de replicar un experimento. A continuación se describen los propósitos de estas formas de replicación.

- *Repetición.* En la repetición se verifica que el mismo investigador en el mismo sitio siguiendo el mismo protocolo pueda reproducir de nuevo los resultados obtenidos en el experimento base. Al reproducir los resultados la probabilidad de obtener un error de tipo I^2 disminuye. Es necesario contar con varias repeticiones de un experimento para conocer la variabilidad natural de los resultados observados. La repetición permite estimar cuánto de igual debieran ser los resultados de posteriores repeticiones del experimento.
- *Replicación.* En la replicación se verifica que los resultados experimentales son 1) suficientemente objetivos para ser observados por otros investigadores, 2) suficientemente globales para ser observados en otros laboratorios y 3) suficientemente estables para ser observados con cambios en el protocolo experimental empleado. Una replicación utiliza el mismo protocolo experimental al del experimento base. Diferentes tipos de repeticiones son necesarias para verificar la independencia de resultados de los diferentes elementos que conforman la estructura de un experimento.
- *Reproducción.* En la reproducción se verifica que los resultados experimentales no son atribuibles al protocolo experimental. En una reproducción se lleva a cabo un nuevo experimento con un protocolo experimental distinto para comprobar las mismas hipótesis del experimento base. En una reproducción, el investigador parte del conocimiento de un hecho empírico el cual ha sido previamente establecido por otro investigador.

4.2 La Repetición como forma de Replicación en ISE

En un experimento controlado el tamaño de la muestra representa un elemento de vital importancia. El contar con un tamaño de muestra adecuado aumenta las posibilidades de que el efecto observado en la muestra se da en la población. El tamaño de la muestra influye en la precisión con la que se puede conocer el efecto que se estudia. Cuanta mayor precisión se busque, el tamaño de la muestra debe ser mayor.

Uno de los coeficientes comúnmente empleado para representar el tamaño del efecto observado en un experimento es el coeficiente d (Cohen, 1988). Este coeficiente se usa para medir las diferencias entre dos variables, en este caso dos de los tratamientos estudiados en

²Se comete cuando se rechaza la hipótesis nula y ésta es verdadera i.e. cuando se cree que hay diferencia significativa entre los tratamientos comparados en el experimento y en realidad no la hay.

el experimento. El tamaño del efecto permite conocer qué tanto es mejor un tratamiento frente a otro. Este coeficiente suele utilizarse con una precisión de un dígito. Por ejemplo (Cohen, 1988), $d=0.2$ representa un efecto pequeño, $d=0.5$ indica un efecto mediano o $d=0.8$ es un efecto grande. Si un dígito es el nivel de precisión requerido para obtener conocimiento acerca del tamaño de un efecto, entonces a partir de (4.1) es posible calcular el tamaño de la muestra necesario para satisfacer el nivel de precisión de un dígito: la función en (4.1) se deriva de (4.2) y (4.3) donde la diferencia de los intervalos de confianza (izquierdo y derecho) se igualan al nivel de precisión que se esté buscando, en nuestro caso 0.1.

$$N = \frac{2 + d^2}{2(0.0255102)^2} \quad (4.1)$$

$$2 \times 1.96 \times \text{desviación}(d) = 0.1 \quad (4.2)$$

$$\text{desviación}(d) = \sqrt{\frac{n1 + n2}{n1n2} + \frac{d^2}{2(n1 + n1)}} \quad (4.3)$$

Para el caso de los tamaños de efecto $d=0.2$, $d=0.5$ y $d=0.8$ se requiere de un tamaño de muestra de $N=1,567$, $N=1,729$, $N=2,028$ respectivamente. En la figura 4.1, se muestra la gráfica de la función resultante en (4.1).

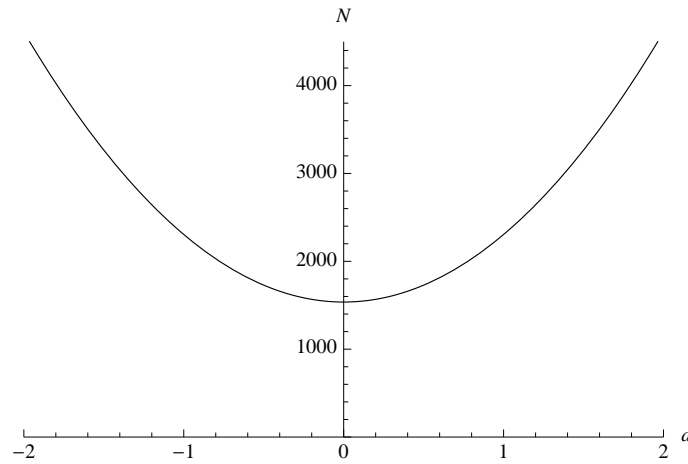


Fig. 4.1: Muestra necesaria para un tamaño de efecto determinado con una precisión de un dígito.

De acuerdo al conjunto de experimentos controlados en IS estudiados por Dybå et al. (Dybå et al., 2006), el tamaño promedio de las muestras usadas en estos experimentos es de $N=55$ (55 observaciones por experimento). En IS son raros los experimentos que llegan a tener un tamaño de muestra mayor a 100 observaciones ((Juristo y Vegas, 2003; Cruz-Lemus et al., 2005; Williams et al., 2003) son ejemplos de algunos de ellos).

Si se desea tener una precisión de un dígito para estimar los tamaños de los efectos en nuestra disciplina, se requieren repeticiones de un mismo experimento que aumenten el tamaño de la muestra. Para un tamaño promedio de muestra de 50 observaciones, se necesita repetir 31 veces el mismo experimento para satisfacer el tamaño de muestra para un efecto de tamaño $d=0.2$; Para el caso de los tamaños de efecto $d=0.5$ y $d=0.8$ se necesita repetir el mismo experimento 34 y 40 veces respectivamente.

Cabe señalar que desde esta perspectiva estadística las repeticiones del experimento tienen que ser iguales i.e. (todas) las replicaciones tienen que medir exactamente igual las variables independientes y dependientes, se tiene que emplear exactamente el mismo protocolo experimental y en todas las replicaciones se muestrean en las mismas poblaciones (Hunter, 2001).

Dado las dificultades de control de las condiciones experimentales en ISE, una opción a considerar para satisfacer la necesidad estadística de repetición es realizar rutinariamente repeticiones (en el mismo sitio y por los mismos investigadores) de un mismo experimento.

Este requisito no resulta extravagante en disciplinas experimentalmente más maduras, como psicología donde suelen usar tamaños de muestra entre 100 y 200 observaciones, incluso 500 observaciones (Wasserman y Bracken, 2003).

Mediante repeticiones como forma de replicación, la muestra se acerca más al intervalo de observaciones $[1,537; 2,305]$ necesario para tener confianza que el efecto observado (desde 0, ninguno, a 1, muy grande) se produce no sólo en la muestra empleada en el experimento sino también en la población. En este caso la población local que está muestreando este experimento.

Establecer un hecho a partir de los resultados de una sola ejecución de un experimento implica una mayor probabilidad de que el resultado esté amenazado por el error tipo I³. El contar con más repeticiones del mismo experimento reduce considerablemente este tipo de error. Por ejemplo, si un investigador establece en un experimento el nivel de significación α en 0.05 que representa la probabilidad de 1:20 de obtener un resultado fortuito, al repetir de nuevo este experimento la probabilidad de obtener de nuevo un resultado fortuito disminuye a 1:400 ($p = 0.05 \times 0.05 = 0.0025$).

Los experimentos realizados en IS carecen del tamaño de muestra para estimar de manera precisa el tamaño del efecto estudiado. Por tanto son necesarias repeticiones como forma de replicación para poder estimar el tamaño del efecto con cierta precisión. Sin embargo las repeticiones son prácticamente imposibles en ISE cuando se realizan en otros sitios (Juristo y Vegas, 2009).

Cuando una disciplina experimental es joven existen muchos aspectos que se desconocen de ésta, cualquier cambio en una replicación por mínimo que sea puede ocasionar una variación no entendida en los resultados del experimento (Collins, 1985). Es en este tipo de disciplinas como la ISE donde realizar repeticiones de un mismo experimento ayuda a conocer la variación natural de los resultados observados así como reducir la probabilidad de cometer un error de tipo I. En este tipo de disciplinas inmaduras comenzar a realizar repeticiones de un mismo experimento ayuda al investigador a conocer las condiciones experimentales que deben controlarse.

³El error tipo I consiste en rechazar la hipótesis nula cuando es verdadera i.e. cuando se cree que hay diferencia significativa entre los tratamientos comparados en el experimento y en realidad no la hay.

4.3 Replicaciones Independientes

Otro concepto que se ha malinterpretado en ISE es el relacionado con las replicaciones independientes. Cuando se piensa en la replicación se tiene como estereotipo de replicación aquella que debe llevarse a cabo de forma independiente por otros investigadores en distintos sitios a los del experimento base, por lo que parecería implicar que en la replicación no hay comunicación entre el replicador y el investigador del experimento original.

No obstante, se observa que al realizar replicaciones independientes existe cierta comunicación entre replicadores e investigadores. Por ejemplo, en el experimento sobre fusión fría realizado por Fleischmann y Pons (Fleischmann y Pons, 1989) hubo interacción entre los autores del experimento y otros investigadores quienes intentaron realizar replicaciones de este experimento (Browne, 1989).

Otro ejemplo de comunicación entre replicadores e investigadores del experimento original se encuentra en el número 7101 de Nature (Giles, 2006). En uno de los artículos Jiang et al. (Jiang et al., 2002) reportan haber logrado aislar completamente células madre de una médula humana adulta. Tras esta publicación, varios grupos de investigación contactaron a los autores para pedir más detalles sobre el protocolo experimental usado, inclusive varios de estos grupos enviaron investigadores al sitio donde se realizó el experimento original para recibir adiestramiento en el protocolo experimental usado.

Un ejemplo más, se encuentra estudiado en profundidad en el libro de Collins (Collins, 1985) en donde diferentes laboratorios realizaron replicaciones para construir un láser de descarga transversal eléctrica a presión atmosférica o láser TEA CO₂ (En inglés, Transversely-Excited Atmospheric-pressure CO₂ laser).

El primer láser de este tipo fue construido en 1968 por un grupo de investigación canadiense, donde en principio los resultados se mantuvieron en secreto. Después del anuncio de la construcción del láser TEA CO₂ (Beaulieu, 1970), otros laboratorios intentaron construir copias de este aparato (Gibson et al., 1971; Patel, 1973; Biswas y Harrison, 1986). Collins (Collins, 1985) visitó y habló con investigadores de varios de estos laboratorios y encontró que:

- “ningún científico tuvo éxito en la construcción de un láser usando solamente la información encontrada en fuentes publicadas u otras fuentes escritas.”
- “ningún científico tuvo éxito en construir una láser TEA CO₂ donde su informante fuese un ‘hombre promedio’ quién no ha construido un dispositivo por sí mismo.”
- “... incluso cuando el informante había construido con éxito el dispositivo, ... , el principiante sería poco proclive a tener éxito sin un cierto período de contacto con el informante y, en ciertos casos, no tendría éxito en absoluto.”

De acuerdo a lo encontrado en otras disciplinas, es aceptable en ISE realizar la replicación de forma independiente manteniendo cierto grado de comunicación entre el replicador y el investigador quien realizó el experimento base.

4.4 Elementos de una Situación Experimental

El nivel de similitud entre una replicación y el experimento base varía dependiendo de los cambios realizados en la replicación. Para conocer los diferentes cambios que puede haber en la replicación es necesario identificar las partes o elementos que conforman un experimento o situación experimental típica en ISE.

Se han identificado cinco secciones en una situación experimental típica de ISE que son: experimentadores, sitio, protocolo experimental, operacionalizaciones de los constructos y propiedades de las poblaciones. A continuación se describe cada uno de los elementos que conforman las distintas secciones de la situación experimental.

4.4.1 Experimentadores

Durante la realización de un experimento los investigadores pueden desempeñar más de un rol. Por ejemplo, una persona puede participar en el diseño del experimento, otra durante la ejecución y otra más en el análisis. Se proponen cinco roles que pueden variar en un experimento, éstos son: diseñador, entrenador, monitor, medición y analista.

- El diseñador se encarga de: definir y especificar la estructura de los tratamientos que se asignan a los sujetos; Definir la operación del experimento y planificar la ejecución de éste; Así mismo elabora estrategias para aumentar el control sobre posibles variables extrañas mitigando las amenazas a la validez del experimento.
- El entrenador es responsable de transmitir el tratamiento a los sujetos durante la sesión de entrenamiento, el tratamiento puede ser transmitido durante una asignatura, curso, seminario o taller.
- El monitor cuida a los sujetos durante la sesión del experimento, sus funciones son dar algunas instrucciones a los sujetos referentes a cómo se llevará a cabo la sesión, entregar a los sujetos los materiales que incluyen: tratamientos, formularios, objetos experimentales y herramientas, recoger los materiales al final de la sesión, y aclarar dudas. Nótese que la aclaración de dudas puede alterar el concepto de los tratamientos transmitidos a los sujetos al inicio de la sesión.
- El rol de medición realiza el procedimiento de medición sobre los datos recabados durante el experimento.
- El analista efectúa los análisis estadísticos, sus funciones incluyen la identificación de posibles valores atípicos (*outliers*), elaboración de análisis descriptivos, pruebas de significación, pruebas de comparación post-hoc, entre otros análisis.

En una replicación es conveniente que los investigadores varíen los roles para asegurar la independencia de resultados, al variar los roles se evita que los investigadores mantengan alguna preferencia o prejuicio hacia los resultados.

4.4.2 Sitio

Este elemento identifica el lugar donde se realiza la replicación. Cuando se obtiene cierto resultado de un experimento, se espera que éste pueda reproducirse en otro sitio incluso replicando lo más exacto posible el experimento. Al reproducir los resultados en otro sitio se asegura que el resultado observado no es local producto del sitio donde se realizó el experimento base.

4.4.3 Protocolo experimental

El protocolo experimental se refiere a los instrumentos, materiales, objetos experimentales, formularios y procedimientos empleados en un experimento. En este sentido el protocolo es la configuración de todos estos elementos que permite al investigador observar cierto efecto o evento (resultados experimentales). Los elementos del protocolo

experimental identificados que pueden variar en una replicación son: diseño experimental, técnicas de análisis de datos, instrumentos o equipos de medición, objetos experimentales, formularios, recursos necesarios para aplicar los tratamientos así como guías suplementarias.

- **Diseño experimental.** El diseño experimental dicta la forma en cómo se deben organizar los tratamientos que se asignan a los sujetos. Existen diversos tipos de diseños experimentales que, dependiendo del tipo de diseño que se use, es posible examinar interacciones entre los tratamientos y otras condiciones o elementos de la situación experimental. Por ejemplo, si se desea comparar la efectividad de dos técnicas de inspección de código (tratamientos) en sujetos con una experiencia *junior* y *senior*, se usa un diseño experimental factorial 2^2 con dos factores (técnica y experiencia) y dos niveles por factor (Tec. A, Tec. B, *junior* y *senior*) donde además de conocer la efectividad de ambas técnicas se examina una posible interacción de la experiencia de los sujetos en las técnicas.
- **Técnicas de análisis de datos.** Este elemento se refiere a la forma en cómo se analiza estadísticamente el conjunto de observaciones recabadas tras ejecutar el experimento. Por ejemplo, para el tipo de diseño factorial 2^2 descrito anteriormente, se suele utilizar un modelo estadístico conocido como ANOVA (*Analysis of Variance*, Análisis de la varianza) donde se examinan los resultados de los tratamientos para determinar si existe alguna diferencia significativa i.e. si alguno de los tratamientos se comporta mejor que otro.
- **Instrumentos o equipos de medición.** Este elemento se refiere a los instrumentos o equipos de medición que pueden emplearse para medir la variable de respuesta del experimento. Por ejemplo, una variable de respuesta relacionada con el tiempo de duración de alguna tarea puede medirse de forma manual (el sujeto lleva un registro manual del tiempo empleado en alguna tarea) o automática se dispone de un instrumento que registra automáticamente el tiempo de la tarea.
- **Objetos experimentales.** Los objetos experimentales son los artefactos tales como especificaciones, diagramas, códigos fuente o programas que se entregan a los sujetos durante el experimento para que completen una serie de tareas relacionadas con algún aspecto de la IS.
- **Formularios.** Los formularios se utilizan para recabar información de las tareas que realizan los sujetos durante la sesión del experimento.
- **Guías suplementarias.** Las guías suplementarias orientan al sujeto sobre algún aspecto relacionado con el tratamiento o con el objeto experimental. Por ejemplo, en un experimento donde se compara la efectividad de varias técnicas de inspección de código, las guías suplementarias pueden contener información referente a alguna de las funciones o procedimientos implementados en los códigos fuente.

4.4.4 Operacionalizaciones de los Constructos

Las operacionalizaciones representan el acto de traducir un constructo en su manifestación. En un experimento controlado se tienen operacionalizaciones de causa y efecto. Ambos tipos de operacionalizaciones contienen elementos que pueden variar en una replicación.

Respecto a las operacionalizaciones de causa se identifican los siguientes elementos que pueden estudiarse en las replicaciones:

- Tratamiento acorde a la literatura. El investigador utiliza alguna fuente (e.g. la literatura) para seleccionar los tratamientos (operacionalizaciones de causa) que desea examinar.
- Procedimiento de aplicación del tratamiento. El investigador adapta de la literatura a su experimento el procedimiento de aplicación de los tratamientos.
- Tratamiento como instrucciones. Una vez adaptados los tratamientos, éstos son transformados en una serie de instrucciones que el sujeto debe realizar durante las sesiones del experimento.
- Recursos del tratamiento. La aplicación del tratamiento puede requerir de algún tipo de software o hardware. Dependiendo del experimento, los recursos pueden ser manuales i.e. no se requiere el uso de ordenador; o automatizados, como puede ser algún tipo de herramienta software, sistema operativo o tipo de ordenador.
- Aspectos de transmisión del tratamiento. La forma en cómo se transmiten los tratamientos a los sujetos puede influir en los resultados. Por ejemplo, los resultados pueden influir si el investigador responsable de transmitir el tratamiento a los sujetos no cuenta con suficiente experiencia en los tratamientos que se utilizan en el experimento.

Como se observa, en las operacionalizaciones de causa se identifican dos niveles de abstracción, en el primer nivel el investigador adapta de alguna fuente (e.g. la literatura) a su contexto el procedimiento de aplicación de los tratamientos (operacionalizaciones de causa). Una vez adaptados los tratamientos se transforman en una serie de instrucciones que el sujeto debe realizar durante las sesiones del experimento.

Por otra parte, en las operacionalizaciones de efecto se identifican los siguientes elementos que pueden variar en las replicaciones:

- Métrica. La métrica representa la operacionalización de efecto que el investigador define para medir el constructo de efecto. Por ejemplo, para medir la efectividad (constructo) de una técnica de inspección de código, el investigador puede operacionalizar este constructo a través de la métrica “porcentaje de faltas observadas”.
- Procedimiento de medición. El investigador especifica el procedimiento de medición en el que se describe como calcular y asignar a la variable de respuesta las observaciones obtenidas. Las variables de respuesta se derivan de las métricas previamente definidas. Este procedimiento puede ser sencillo, por ejemplo, obtener de forma directa los valores de la variable de respuesta o complejo como puede ser utilizar una serie de formulas para calcular a partir de las observaciones los valores de la variable de respuesta. En algunos casos este procedimiento puede automatizarse de tal forma que las observaciones recabadas durante el experimento se asignen directamente a las variables de respuesta.

4.4.5 Propiedades de las Poblaciones

En experimentos controlados en IS, se identifican dos tipos de poblaciones que son de interés poder generalizar, el primer tipo de población se relaciona con los sujetos que participan en un experimento. Las propiedades de los sujetos son las que pueden variar en una replicación. El otro tipo de población identificado se refiere a los objetos experimentales con los que trabaja el sujeto durante el experimento. Por objetos experimentales se refiere a las especificaciones, documentos de diseño, códigos fuente, programas u otros tipos de artefactos afines al proceso de desarrollo software.

Dependiendo de los cambios que se realicen en los elementos de la situación experimental se consiguen distintos tipos de comprobaciones respecto a los resultados observados en el experimento base. En la tabla 4.1 se muestra el avance de conocimiento que se obtiene al variar en la replicación los distintos elementos de la situación experimental.

Elemento	Tipo de conocimiento ganado cuando	
	Hay cambios	No hay cambios
Roles de los experimentadores	Evento objetivo	Evento subjetivo
Sitio	Evento global	Evento local
Protocolo experimental	Evento real	Evento artificial
Operacionalizaciones de los constructos	Límites de operacionalizaciones conocidas	Evento limitado a la forma de aplicar los tratamientos y medir las respuestas
Propiedades de las poblaciones	Límites de poblaciones conocidas	Evento limitado a ciertas propiedades de las poblaciones

Tabla 4.1: Conocimiento ganado al variar elementos de la situación experimental

4.5 Funciones de la Replicación en ISE

Cuando se realiza una replicación interesa tanto mantener sin cambios los elementos de la situación experimental para comprobar que el efecto observado es reproducible, como interesa variar los elementos de la situación experimental ya sea para verificar que el efecto observado no es artificial o para conocer los límites bajo los cuales es posible reproducir el efecto observado. Es decir la replicación requiere tanto similitud como dis-similitud respecto al experimento base, dependiendo del tipo de verificación a realizar.

La función principal de la replicación es verificar los resultados de un experimento, no obstante, son necesarios varios tipos de verificación (Juristo y Gómez, 2011). Por ejemplo, replicar lo más similar posible un experimento verifica que los resultados no sean fortuitos, o variar las propiedades de la población del experimento base verifica los límites de la población para los que se cumplen los resultados. La similitud entre el experimento base y la replicación varía dependiendo del objetivo de verificación que busque la replicación. Se han identificado los siguientes objetivos de verificación que una replicación en ISE puede tener.

- *Control del error muestral.* Si los elementos de la situación experimental se mantienen sin cambios, la función de la replicación es verificar que los resultados obtenidos en el experimento base no son fortuitos producto de un error de tipo I.
- *Independencia de los experimentadores.* Si distintos investigadores realizan la replicación, entonces la función de la replicación es verificar que los investigadores no influyen en los resultados.
- *Independencia del sitio.* Si la replicación se lleva a cabo en otro sitio, entonces la función de la replicación es verificar que los resultados son independientes del sitio donde se realiza el experimento.
- *Independencia del protocolo experimental.* Si se varían los elementos del protocolo experimental, entonces la función de la replicación es verificar que los resultados

observados no son artificiales i.e. los resultados son reales y no artificiales producto de la configuración del protocolo experimental.

- *Definir límites en operacionalizaciones.* Si las operacionalizaciones de los constructos se modifican, entonces la función de la replicación es determinar los límites hasta donde se reproducen los resultados al variar los elementos que constituyen estas operacionalizaciones.
- *Definir límites en poblaciones.* Si las propiedades de las poblaciones varían, entonces la función de la replicación es determinar los límites hasta donde los resultados pueden extenderse a otros tipos de sujetos u objetos experimentales.

Si la replicación del experimento base se realiza a través de una reproducción (repetición del experimento que únicamente conserva las hipótesis) entonces es posible observar el mismo resultado con diferentes experimentos de objetivos idénticos.

Todas las funciones son importantes en la replicación ya que ayudan a estudiar los niveles de variación (control del error muestral), garantizar la independencia de resultados con respecto a los investigadores y sitio, evitar la generación de eventos artificiales del protocolo experimental y comprender los límites de las operacionalizaciones y las poblaciones bajo las cuales los resultados del experimento se mantienen.

Estas funciones se relacionan con los diferentes tipos de validez (Cook y Campbell, 1976). Si la replicación se realiza para controlar el error muestral, aumenta la confianza en la validez de conclusión, ya que se verifica que la relación observada es verosímil. Si la replicación se realiza para verificar la independencia de los experimentadores, sitio o protocolo experimental entonces aumenta la confianza en la validez interna, ya que se verifica que otras condiciones no influyen en los resultados del experimento. Si la replicación se realiza para determinar los límites de las operacionalizaciones, entonces aumenta la confianza en la validez de los constructos, ya que se evalúa la traducción de los constructos a los tratamientos primarios y medidas (operacionalizaciones de causa y efecto). Finalmente, si la replicación se realiza para determinar los límites de las propiedades de las poblaciones, aumenta la validez externa, ya que se examina si los resultados son generalizables a otros sujetos u otros objetos experimentales.

En este capítulo se han clarificado conceptos con el fin de aligerar su incompreensión en ISE, en el siguiente capítulo se desarrolla una tipología de replications para la ISE.

Capítulo 5

Tipología de Replicaciones para la ISE

RESUMEN: En este capítulo se desarrolla una tipología de replicaciones para la ISE, así como se describe un enfoque sistemático para realizar replicaciones usando como referencia la tipología propuesta.

5.1 Tipología de Replicaciones para la ISE

De acuerdo a los diferentes tipos de replicación encontrados en otras disciplinas, un tipo de replicación puede componerse de diferentes calificativos. Por ejemplo, de acuerdo a los tipos de replicación *parcial* e *independiente* definidos en (Hendrick, 1990) y en (Lüdtke, 2008), se puede definir una replicación como *parcial-independiente* a aquella replicación que varía diferentes elementos que se creen irrelevantes para los resultados del experimento y que se realiza por investigadores distintos a los del experimento base. En otras palabras, una replicación puede ser calificada con más de un adjetivo cuando tiene diferentes propiedades que la caracterizan.

Los tipos de replicación propuestos en este trabajo de investigación se componen de varias propiedades relacionadas con:

- El nivel de semejanza entre la replicación y el experimento base
- La conexión entre replicadores e investigadores del experimento base
- El sitio donde se realiza la replicación.

De acuerdo a lo anterior, a continuación se presentan los tipos de replicación propuestos para la ISE.

5.1.1 Por Semejanza

Dependiendo de la **semejanza** que hay en los elementos de la situación experimental y del experimento base se proponen los siguientes tipos de replicación:

- **Exacta.** Se realiza una copia idéntica del experimento base. Sin embargo esto no es posible en el estado actual de la ISE pues debido al desconocimiento de las variables relevantes del contexto siempre existen condiciones que varían intencionada o inadvertidamente entre la replicación y el experimento base. En algunas áreas de las ciencias naturales es posible llegar a una aproximación bastante cercana de este tipo de replicación ya que se tiene un mayor control sobre las condiciones. Por ejemplo, en estas áreas se trabaja en laboratorios estériles, libres de variables extrañas que puedan afectar los resultados, es posible utilizar cantidades exactas de compuestos o materiales, así como utilizar ingredientes o componentes que no varían con el tiempo.
- **Literal.** Esta replicación es el equivalente en ISE a la replicación exacta. Se intenta realizar la replicación lo más exacta posible al experimento base. Los elementos de la situación experimental se mantienen igual, es decir, la replicación se realiza siguiendo el mismo protocolo, con las mismas operacionalizaciones y poblaciones equivalentes a las del experimento base. En esta replicación no se realizan cambios deliberados en los elementos de la situación experimental.
- **Artificial.** Se varían elementos del protocolo experimental con el fin de verificar que los resultados observados se reproducen con protocolos experimentales equivalentes. Es decir, los resultados no son artificiales.
- **Operacional.** Se varían las operacionalizaciones de causa y/o efecto con el fin de verificar hasta qué límites de las operacionalizaciones de los constructos de causa y/o efecto se mantienen los resultados. En otras palabras, se estudia la sensibilidad de los resultados con respecto a las operacionalizaciones.
- **Poblacional.** Se varían las poblaciones para verificar los límites de las poblaciones usadas en el experimento base. En otras palabras, se estudian las propiedades de las poblaciones que influyen en los resultados.
- **Conceptual.** En esta replicación el investigador elabora su propio protocolo experimental y operacionalizaciones para verificar los resultados observados en el experimento base.

Las replications artificiales, operacionales, poblacionales y conceptuales pueden identificarse como replications **diferenciadas** con la finalidad de usar un nombre contrapuesto a la replicación literal para denotar una replicación que ha sufrido cambios respecto al experimento base.

5.1.2 Por Conexión entre Experimentadores

Según el nivel de conexión entre los investigadores que realizaron el experimento base y los que realizan la replicación, se proponen los siguientes tipos:

- **Nativa.** Si se realiza por los mismos investigadores que participaron en el experimento base.
- **Conjunta.** Parte de los investigadores que participaron en el experimento base participan en la replicación.
- **Ajena.** Se realiza por investigadores ajenos a quienes participaron en el experimento base.

5.1.3 Por Lugar

Según el lugar donde se realiza. Los tipos de replicación propuestos son:

- **Interna.** Si se realiza en el mismo sitio donde se efectuó el experimento base.
- **Externa.** Si se realiza en otro sitio distinto al del experimento base.

La replicación literal-nativa-interna puede denominarse también como **repetición**. De manera similar la replicación conceptual-ajena-externa puede denominarse también como **reproducción**.

5.2 Cambios Sistemáticos en las Replicaciones

Los tipos de replicación propuestos asociados a las diversas funciones de verificación pretenden ayudar a replicar de forma sistemática. Es decir, realizar cambios conociendo de antemano el tipo de verificación que se alcanzará debido, precisamente, a tales cambios: aumentar la confianza en que el resultado no es artificial, explorar los límites en las operacionalizaciones, explorar los límites en las poblaciones así como explorar variables relevantes de las poblaciones.

El hecho de que ocurran cambios incluye que estos vengan impuestos por el nuevo contexto. Por ejemplo, al realizar una replicación el nuevo contexto puede obligar a cambiar el diseño: el experimento base se realizaba en 3 días y todos los sujetos aplicaban 3 técnicas, mientras que en el nuevo contexto no se cuenta con tanto tiempo y el diseño cambia a un día donde 1/3 de los sujetos aplica cada técnica. Este cambio grande en el experimento lejos de ser una amenaza a la verificación de los resultados puede resultar una oportunidad de comprobar si determinado elemento del protocolo experimental (el diseño) afecta a los resultados.

Si ésta es la primera replicación del experimento y se obtienen resultados diferentes, no se podrá asignar la variación al diseño de modo seguro (pueden estar actuando variables desconocidas que han sido variadas sin querer). Pero a la larga, cuando vayan realizándose más replicaciones, esta replicación jugará su papel en la verificación de resultados. Es más, si el experimento base se hubiera repetido (replicación literal-nativa-interna) varias veces por los investigadores iniciales, se contaría ya con una buena estimación de la variación natural de los resultados y esto ayudaría a que la nueva replicación (artificial-ajena-externa) pueda contribuir antes a la verificación de resultados: quizás lo que pareciera resultados distintos entre la nueva replicación y el experimento base en realidad esté dentro de los límites de la variación natural, ahora que el experimento base cuenta con una muestra mayor de resultados gracias a las repeticiones realizadas por los mismos investigadores en sus propios laboratorios.

Variar de forma sistemática las replicaciones (incluidas las nativa-interna tanto artificial, operacional y poblacional) contribuye a aumentar la comprensión de las condiciones que pueden influir en los resultados de un experimento.

Conforme se conocen las diferentes condiciones que influyen en los resultados de un experimento, las observaciones se vuelven mucho más exactas así como se amplía la verificación de los resultados del experimento.

Variar sistemáticamente las replicaciones no es novedoso. Desde hace más de tres décadas Hunt (Hunt, 1975) sugirió que un mejor procedimiento para realizar replicaciones es utilizar un enfoque sistemático en el que se modifique en cada replicación un elemento del experimento original con el fin de estudiar si ese cambio influye en los resultados.

Como afirma Hendrick (Hendrick, 1990) las replicaciones sistemáticas aplicadas con rigurosidad, pueden cambiar las actitudes negativas hacia las replicaciones. De manera similar, Rosenthal (Rosenthal, 1990) propone realizar baterías de replicaciones donde al menos se requieren dos replicaciones, una lo más similar al experimento base y la otra moderadamente dis-similar.

Seguir un enfoque sistemático permite a los investigadores avanzar de una forma gradual y lógica en el proceso de verificación de resultados experimentales. En IS se desconocen las condiciones que influyen en los experimentos que se realizan. Utilizar un enfoque sistemático donde se realicen replicaciones nativas-internas con cambios deliberados para comprender la dependencia del protocolo experimental, la sensibilidad a las operacionalizaciones y las características relevantes de las poblaciones contribuiría a aumentar el conocimiento sobre condiciones experimentales relevantes (y, de paso, de variables relevantes en el desarrollo de software).

Tras un único experimento no se saben si los resultados observados son producto de una coincidencia. El primer paso hacia la verificación de resultados consiste en realizar repeticiones del experimento, donde el mismo investigador, en el mismo sitio, con el mismo protocolo, con las mismas operacionalizaciones, y con el mismo tipo de poblaciones verifique el resultado observado. La repetición del experimento ayuda a conocer la variación natural de los resultados observados. Es decir, el intervalo de confianza donde se observan los resultados, así como reducir las probabilidades de cometer un error Tipo I.

Tras varias repeticiones se desconoce si los resultados son artificiales o reales. Los eventos observados pueden ser producto de la configuración del protocolo empleado en el experimento. Una vez observada una regularidad (resultados similares en replicaciones literales-nativas-internas), el siguiente paso es verificar que ésta no es producto del protocolo. Procede, ahora, realizar replicaciones donde se varíe el protocolo experimental.

Tras esta serie de replicaciones, los resultados observados de las diferentes replicaciones pueden considerarse independientes del protocolo, es decir, los eventos observados corresponden a alguna relación causal existente en la realidad.

Una vez que obtenemos regularidades reales, es posible variar las poblaciones y las operacionalizaciones para conocer los límites en los que podemos reproducir los resultados del experimento. Al variar estos dos elementos se obtienen nuevos conocimientos sobre características críticas de las poblaciones y sobre los “principios activos” de los factores y variables estudiados.

En este punto el investigador cuenta con varias replicaciones donde ha variado elementos del protocolo experimental así como las propiedades de las poblaciones y las operacionalizaciones. Con los resultados de estas replicaciones el investigador conoce la variación natural regular de los resultados (intervalo de confianza) así como bajo qué condiciones y límites puede reproducir un resultado experimental. Este conocimiento permite entender mejor los resultados producidos por replicaciones externas y ajenas.

Cuando se tiene mayor madurez sobre las condiciones que influyen en los resultados del experimento, es posible realizar una replicación conceptual-ajena-externa (reproducción) con el fin de verificar por distintos caminos los resultados.

La reproducción tiene el mayor poder de confirmar un resultado experimental, ya que el resultado se verifica con distintos experimentadores, en sitios diferentes, con un protocolo diferente, con operacionalizaciones diferentes y con poblaciones equivalentes a las del experimento base. Sin embargo, esta forma de replicación es la más arriesgada ya que si no se obtienen resultados semejantes es difícil (de hecho imposible) encontrar las causas de las diferencias en los resultados entre la replicación y el experimento base.

De ahí que resulte una aproximación más barata y razonable comenzar con variaciones pequeñas y llegar a la máxima variación al final del proceso.

En la figura 5.1 se presenta un orden de replicaciones que minimiza el riesgo de interpretación conjunta de los resultados de replicaciones y experimento base. El eje y indica el nivel de potencia de confirmación de los resultados experimentales mientras que el eje x muestra el riesgo que implica llevar a cabo las diferentes formas de replicación. Las primeras replicaciones (repeticiones) se realizan de forma nativa e interna, en la replicación se asume que puede haber una combinación de tipos entre nativa, conjunta, ajena e interna y externa, no obstante la replicación conceptual (reproducción) se realiza de forma ajena y externa.

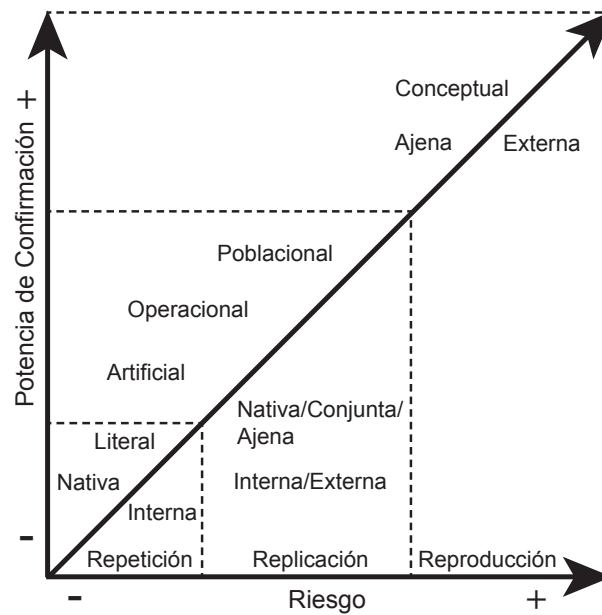


Fig. 5.1: Cambios sistemáticos para realizar replicaciones

En este capítulo se ha abordado el segundo objetivo de esta investigación referente al desarrollo de una tipología de replicaciones para la ISE. En Los siguiente capítulo se presenta un marco de referencia con las ideas clave para llevar a cabo comparaciones de replicaciones y obtener algún conocimiento de ellas.

Capítulo 6

Marco Conceptual para la Comparación de Replicaciones en ISE

RESUMEN: En este capítulo se presentan las ideas clave que conforman el marco conceptual propuesto en esta tesis para lograr comparar distintos tipos de replications y sortear alguno de los problemas que hasta ahora ha habido en ISE para combinar replications.

Como se describió en el capítulo 4, dependiendo de los cambios hechos en una replicación (cambios en la situación experimental), se obtienen diferentes grados de verificación de resultados. Por ejemplo, en una replicación donde se ha variado el protocolo experimental y se han reproducido los resultados del experimento base, se verifica que los resultados son reales y no artificiales. En otras palabras, se verifica que los resultados no son producto del protocolo experimental empleado. De manera análoga, cambios en los diversos elementos de la situación experimental producen verificaciones de distinto tipo.

En la actualidad, los intentos de combinar replications en ISE para extraer algún tipo de conocimiento no han sido fructíferos. Por ejemplo, en los casos donde es posible combinar algún conjunto de replications porque sus resultados se asemejan entre sí, no se discute el grado de verificación alcanzado, simplemente se agregan los resultados. ¿Son los resultados observados independientes de los investigadores o del sitio donde se realizaron las replications? ¿Son los resultados reales o artificiales producto de la configuración del protocolo experimental empleado?. Por otra parte, en los casos donde no es posible combinar un conjunto de replications debido a las diferencias de resultados, no se es capaz de extraer algún tipo de conclusión (por ejemplo, que los resultados son artificiales) e ingenuamente se cree que debido a las diferencias de resultados y a los cambios hechos en las replications, éstas no son comparables. Resultando, por tanto, las replications inútiles pues no contribuyen a obtener ningún tipo de conocimiento.

En este capítulo se presentan las ideas clave que permiten comparar conjuntos de replications y obtener conocimiento tanto cuando los resultados son semejantes, como cuando son diferentes. Las ideas clave que esta investigación propone para lograr extraer conocimiento de replications en ISE son:

- Unidades comparables entre replications
- Cambios y nivel de similitud entre situaciones experimentales

- Variabilidad de los tratamientos
- Grado de verificación logrado
- Identificación de posibles causas de resultados no reproducibles

6.1 Unidades Comparables entre Replicaciones

En una situación experimental existen dos dominios relacionados entre sí, el dominio donde se encuentran las ideas y el dominio donde se concretan estas ideas. En el dominio de las ideas habitan los constructos que pueden ser teorías, intuiciones o hipótesis que tiene el investigador respecto a las tecnologías, metodologías, procesos, técnicas o herramientas de la IS. En este dominio, el investigador supone o intuye una relación de causa-efecto a nivel de constructos. Por ejemplo, el investigador puede intuir que la efectividad (constructo de efecto) de dos técnicas de detección de defectos (constructos de causa) es distinta.

En el segundo dominio el investigador traduce sus teorías, intuiciones o hipótesis en tratamientos y observaciones de un experimento específico; esto es, realiza una operacionalización de los constructos. Por ejemplo, el investigador interesado en comparar la efectividad (constructo de efecto) de dos técnicas de detección de defectos (constructos de causa) operacionaliza estos constructos de tal manera que la efectividad la traduce como faltas detectadas por sujeto (operacionalización de efecto) y las técnicas de detección de defectos las instancia como técnica de partición de clases de equivalencia y técnica por cobertura de decisiones (operacionalizaciones de causa). En la figura 6.1 se muestra la relación existente entre constructos y operacionalizaciones.

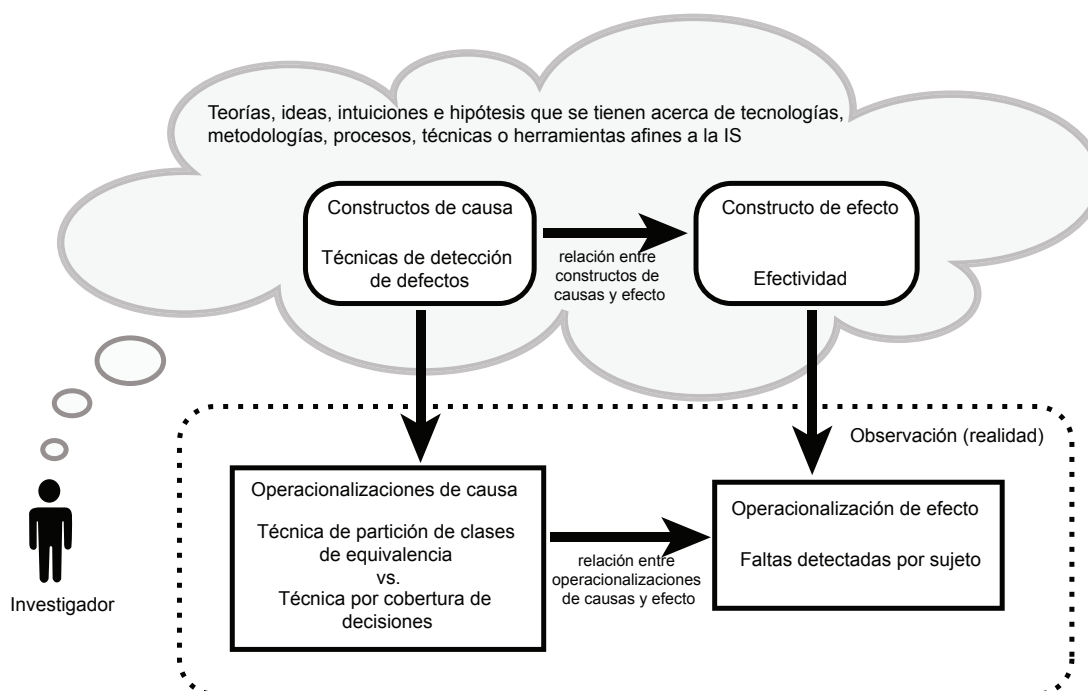


Fig. 6.1: Relación entre constructos y operacionalizaciones (adaptado de Trochim (Trochim, 2000))

Los constructos de causa se conocen también como tratamientos o intervenciones; en el caso de los constructos de efecto, en nuestra disciplina suelen describirse como atributos que pueden ser: productividad, coste, funcionalidad, efectividad, eficiencia, facilidad de uso (*usability*), entre otros más¹. Un constructo no puede medirse de forma directa. Por ejemplo, para medir la facilidad de uso de un sistema es necesario definir uno o varios tipos de medidas que representen este atributo. En este caso las medidas representan las operacionalizaciones del constructo facilidad de uso. Por otra parte, las operacionalizaciones de causa se conocen también como tratamientos o variables independientes, mientras que las operacionalizaciones de efecto se denominan también como variables de respuesta, variables dependientes, medidas o métricas.

De acuerdo a lo anterior, la parte fundamental de una situación experimental son los constructos, las operacionalizaciones y las relaciones de causa-efecto entre éstos. Sin estos elementos, la situación experimental carece de sentido. Sin constructos se desconectan las ideas de sus operacionalizaciones que se encuentran en el mundo real. De manera similar, sin operacionalizaciones se desconectan las observaciones de sus constructos teóricos. Sin relaciones de causa-efecto se desconoce si las ideas a operacionalizar provocan determinado comportamiento.

Para que dos replicaciones sean comparables, éstas deben tener en común al menos dos constructos de causa que están siendo comparados en el experimento. Esto implica que las replicaciones también pueden compararse parcialmente aunque no compartan algunos constructos o que cuenten con un número distinto de constructos. Por ejemplo, una replicación donde se evalúa la efectividad de las técnicas A y B es comparable con otra replicación que evalúa la efectividad de estas mismas técnicas, siendo irrelevante en la comparación, que en la segunda replicación se evalúe también una tercera técnica C. En este sentido el binomio de técnicas A y C, así como B y C de la segunda replicación no pueden compararse con el binomio de técnicas A y B de la primera replicación porque la técnica C no existe en la primera replicación. Una tercera replicación con técnicas C y D no es comparable con la primera replicación con técnicas A y B porque ambas replicaciones tienen pares de constructos de causa disjuntos.

Para ser comparables las replicaciones deben tener en común al menos un constructo de efecto. Por ejemplo, una cuarta replicación que estudia las técnicas A y B no es comparable con la primera replicación si en la cuarta replicación se estudian distintos efectos. Si en la primera replicación se estudia la efectividad mientras que en la cuarta se mide la eficiencia, los resultados de las replicaciones no son comparables. Esto es así porque los resultados de efectividad y de eficiencia no tienen porque corresponderse: puede presentarse la situación donde la técnica A de la primera replicación sea más efectiva que B empero menos eficiente de acuerdo a los resultados de la cuarta replicación. Es decir, las replicaciones primera y cuarta aportan conocimiento complementario, pues no trabajan sobre la misma unidad de conocimiento.

En la figura 6.2 se muestra un ejemplo de las unidades comparables de las cuatro replicaciones usadas como ejemplo. Los recuadros sombreados indican los constructos de las replicaciones que son comparables.

Al despojar la segunda replicación de la técnica C ésta es comparable con la primera replicación ya que ambas replicaciones estudian la misma unidad de conocimiento: los mismos constructos de causa (técnica A y B) respecto al mismo constructo de efecto (efectividad). Como se observa en este ejemplo, **para poder comparar resultados, las**

¹Por ejemplo, en ISO 9126 (ISO, 1999) se describen varios de estos atributos que son utilizados en nuestra disciplina.

Replicación	Causa (Constructos)		Efecto
r1	Técnica A	Técnica B	Efectividad
r2	Técnica A	Técnica B	Efectividad
r3		Técnica C	Efectividad
r4	Técnica A	Técnica B	Eficiencia

Fig. 6.2: Ejemplo de unidades comparables en varias replications

replicaciones deben tener en común dos constructos de causa y un constructo de efecto.

Las replications son comparables aunque los constructos sean operacionalizados de diferentes maneras. Lo necesario para que dos replications sean equivalentes y, por tanto, sus resultados comparables es que compartan una pareja de constructos de causa que sean evaluados con respecto al mismo constructo de efecto. Las operacionalizaciones de los constructos no tienen que ser necesariamente los mismos. Por ejemplo, la facilidad de uso de un sistema software (constructo de efecto) puede medirse a través de un tipo de medida que indique el tiempo que tarda un usuario en completar una tarea (operacionalización de efecto A) o también puede medirse por el número de errores que comete el usuario al completar una tarea (operacionalización B). En este caso el constructo de efecto se concreta con dos operacionalizaciones distintas empero equivalentes.

Por tanto, replications que compartan un mismo constructo de efecto pero distintas operacionalizaciones del mismo pueden compararse siempre que usen operacionalizaciones equivalentes. En otras palabras, si emplean métricas distintas que midan el mismo constructo de efecto. Ciertamente, al comparar los resultados, si éstos no se reproducen, puede ganarse el conocimiento (antes inexistente) de que lo que se consideraba operacionalizaciones equivalentes (por ejemplo métricas equivalentes) en realidad no lo son.

Un constructo de causa puede concretarse con operacionalizaciones diversas. Por ejemplo, una técnica de inspección de código determinada puede llevarse a cabo siguiendo estrictamente el procedimiento descrito por cierta fuente, o se puede variar en distintos grados el procedimiento en aspectos no considerados vitales para aplicar la técnica. En esta situación las replications con operacionalizaciones variadas sirven para averiguar si se mantienen los mismos resultados o si las variaciones en las operacionalizaciones influyen en los resultados. Dependiendo de si los resultados se logran reproducir o no, se ganará el conocimiento de si el resultado no es sensible, o sí, a los cambios en la operacionalización. Por tanto, se revelará que lo que se consideraba variaciones menores, en realidad, afectan al principio activo de la técnica.

Respecto a las operacionalizaciones, las replications **son comparables tanto con operacionalizaciones distintas de un mismo constructo como con variaciones en las operacionalizaciones de los constructos**. Aunque estos cambios deberán tenerse en cuenta a la hora de interpretar los resultados.

De acuerdo a todo lo anterior, para comparar dos o más replicaciones:

- Es necesario que las replicaciones tengan en común al menos un par de constructos de causa, aunque éstas pueden tener distintas operacionalizaciones.
- Es necesario que las replicaciones tengan en común al menos un constructo de efecto. Donde sus operacionalizaciones pueden ser iguales o equivalentes.

En base a estos criterios se propone usar el concepto de partición para denotar la agrupación de unidades comparables entre replicaciones. **La unidad comparable entre replicaciones se conforma con un constructo y operacionalización de efecto y un par de constructos y operacionalizaciones de causa.** Por ejemplo una replicación que evalúa la efectividad de las técnicas A, B y C se puede descomponer en tres unidades conformadas por la evaluación de la efectividad de los pares de técnicas A-B, A-C y B-C. Ciertos pares pueden ser comparables con ciertas replicaciones mientras que otros pares lo serían con otras replicaciones.

Si una segunda replicación evalúa sólo la efectividad de las técnicas A y B, ésta sólo tiene una unidad comparable que es la evaluación de la efectividad de las técnicas A y B. Estas dos replicaciones son comparables porque comparten en común la evaluación de la efectividad de las técnicas A y B. Al agrupar las unidades comparables comunes de ambas replicaciones se obtiene una partición.

Por tanto, **una partición es la agrupación de unidades comparables comunes entre replicaciones.** Si se disponen de varias replicaciones, probablemente se deban agrupar en distintas particiones donde cada partición se conforma por un conjunto de unidades comparables común entre dos o más replicaciones. En la figura 6.3 se muestra un ejemplo donde tres replicaciones se descomponen en unidades comparables agrupadas en dos particiones.

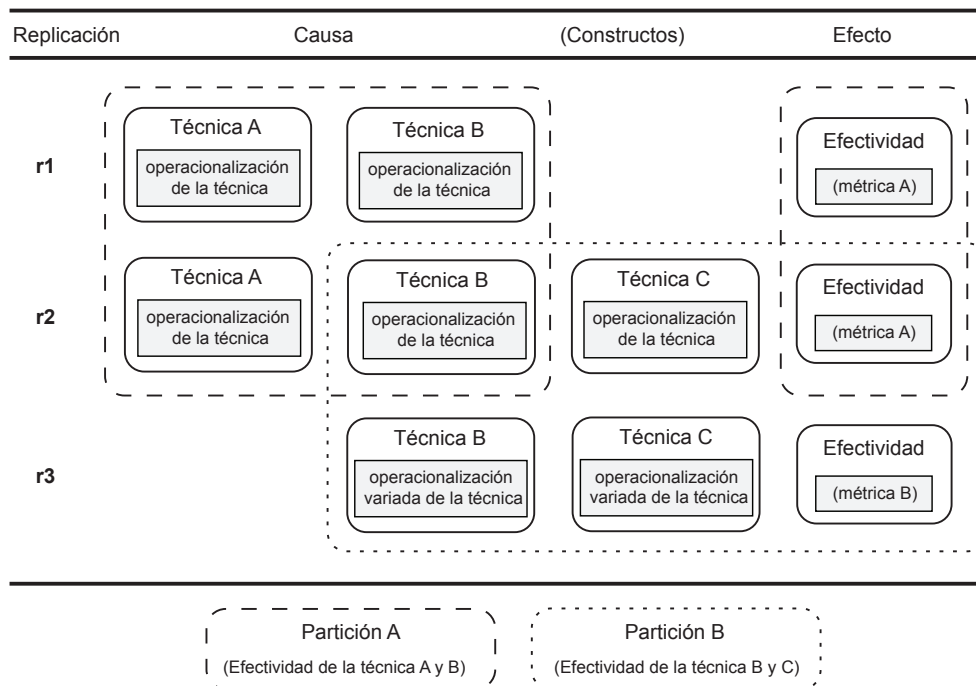


Fig. 6.3: Dos particiones compuestas de varias unidades comparables entre tres replicaciones

En el caso de la partición B, compuesta por la unidad comparable (efectividad de las técnicas B y C) común entre la segunda y tercera replicación, hay diferencias en las operacionalizaciones del efecto (métrica A y métrica B) empero son comparables debido a que las operacionalizaciones pertenecen a un mismo constructo. En $r3$ se ha variado la operacionalización de las técnicas B y C respecto a la segunda replicación $r2$, además, en $r3$ se ha empleado una medida de efectividad distinta (métrica B) a la usada en la segunda replicación (métrica A).

Identificar las unidades comparables entre un conjunto de replicaciones y agruparlas en particiones, permite conocer los grupos de situaciones experimentales y resultados sobre los que se debe estudiar el grado de verificación de resultados. Si el grupo de replicaciones es grande, la estrategia aquí propuesta permitirá descomponer el problema (comparación de replicaciones y resultados) en subproblemas accesibles (comparación de unidades comparables) y atacar su resolución secuencial e iterativamente (una partición tras otra). En el capítulo 7 se propone un procedimiento que el investigador interesado podrá usar para comparar replicaciones y resultados. De manera particular en el capítulo 8 que es parte de este procedimiento, se describe la utilidad y uso de los conceptos de unidades comparables y particiones.

6.2 Cambios y Nivel de Similitud entre Situaciones Experimentales

Cuando se conocen las unidades comparables de un conjunto de replicaciones y se agrupan en particiones, en principio es posible usar alguna técnica de agregación de resultados. No obstante, si se aplica sólo técnicas de agregación de resultados no se identificará el grado de verificación alcanzado cuando los resultados se reproducen, o en caso contrario, será difícil identificar posibles condiciones experimentales que ocasionan resultados no reproducibles en la partición.

Para conocer el grado de verificación alcanzado en un conjunto de replicaciones con resultados reproducidos, **es necesario identificar los cambios hechos en los elementos de la situación experimental de las replicaciones**. Cuando se obtienen resultados contradictorios, identificar cambios hechos en las replicaciones ayuda a conocer posibles condiciones experimentales que parecen influir en los resultados.

Debido, por una parte, a la cantidad de elementos de la situación experimental que pueden variar en una replicación y, por otra, a que distintas replicaciones pueden tener distintos cambios, es necesario registrar minuciosamente los distintos valores tomados por la situación experimental de las replicaciones pertenecientes a una partición.

Nótese que la situación experimental se compone de cinco elementos principales y varios sub-elementos, donde al menos hay 19 elementos particulares que pueden variar en la replicación,

- Cinco afines a los experimentadores (diseñador, entrenador, monitor, medición, analista)
- El sitio donde se realiza la replicación
- Seis afines al protocolo experimental (diseño experimental, técnicas de análisis de datos, instrumentos o equipos de medición, objetos experimentales, formularios, guías suplementarias)
- Siete referentes a las operacionalizaciones (tratamiento acorde a la literatura, pro-

cedimiento de aplicación del tratamiento, tratamiento como instrucciones, recursos del tratamiento, aspectos de transmisión del tratamiento, métrica y procedimiento de medición)

El resto de elementos relacionados con las propiedades de las poblaciones varía en función de los sujetos y objetos experimentales usados en un experimento de IS, por lo que hay más de 19 elementos en una situación experimental.

Al haber en una partición distintas replicaciones, la información de los cambios realizados en los elementos de las situaciones experimentales aumenta de tal forma que es imposible tratar estos cambios mentalmente, se hace necesaria una ayuda donde se registre toda la información para que pueda ser fácilmente comparable.

El registro permite mantener la identificación de cambios, lo que resulta imprescindible para realizar comparaciones entre las situaciones experimentales. En caso de reproducirse los resultados, identificar los cambios asumidos permite conocer el grado de verificación logrado en las replicaciones de la partición, así como condiciones sin influencia en las variables de respuesta. En caso contrario los cambios permiten conocer condiciones que pueden influir en los resultados y que, por tanto, merecen ser investigadas en replicaciones posteriores.

La tabla 6.1 muestra un ejemplo del registro de cambios de los elementos referentes al protocolo experimental de la situación experimental en tres replicaciones. Nótese que el protocolo es sólo uno de los elementos principales de la situación experimental. Lo que significa que la tabla 6.1 es sólo una visión parcial del registro de una situación experimental.

Elemento Principal	Sub Elemento	Protocolo Experimental		
		r1	r2	r3
Análisis	Técnica de análisis de datos	ANOVA 1 vía	ANOVA mixto	ANOVA 1 vía
Diseño	Diseño Experimental	Diseño Aleatorio	Diseño Cruzado	Diseño Aleatorio
	Orden de aplicación de los tratamientos	Aleatorio	2 tipos de orden	Aleatorio
	Combinación del entrenamiento con las sesiones	Secuencial	Secuencial	Secuencial
	Tratamientos por sesión	3 técnicas, 1 programa	2 técnicas, 1 programa	2 técnicas, 1 programa
	Sesiones requeridas por sujeto para aplicar tratamiento	1 sesión	1 sesión	1 sesión
	Tareas por sujeto por sesión	1 técnica, 1 programa	2 técnicas, 1 programa	1 técnica, 1 programa
Instrumentos	Instrumentos o equipos de medición	Ordenador, folios, bolígrafo	Ordenador, folios, bolígrafo	Ordenador, folios, bolígrafo
Objetos experimentales	Programas (Códigos fuente)	Programa A	Programa A	Programa A
	Faltas (instancias)	Instancia de faltas A	Instancia de faltas A	Instancia de faltas A
Materiales	Formularios del experimento	Formularios v1	Formularios v1	Formularios v1
	Instrucciones del experimento	Instrucciones v1	Instrucciones v1	Instrucciones v1
	Guías suplementarias	Guías v1	Guías v1	Guías v1

Tabla 6.1: Registro de cambios del protocolo experimental en tres replicaciones

Los cambios identificados en las situaciones experimentales permiten dar una idea sobre las semejanzas que hay entre las replicaciones de una partición i.e. las semejan-

zas entre las unidades comparables comunes en las replications donde cada unidad se conforma por un constructo y operacionalización de efecto así como por un par de constructos y operacionalizaciones de causa. Como se describe en el capítulo 4, la situación experimental se conforma por cinco secciones o tipos de elementos principales que a su vez contienen elementos específicos que pueden variar en la replicación.

Dependiendo de los cambios realizados en la situación experimental se obtienen distintos tipos de verificación. En la tabla 6.2 se muestran los tipos de verificación obtenidos al variar los elementos de la situación experimental. Estos tipos de verificación se traducen también como tipo de conocimiento ganado. Los elementos principales de la situación experimental se comprueban variando sus sub-elementos, en caso de haberlos.

Elemento	Tipo de conocimiento ganado cuando	
	Hay cambios	No hay cambios
Roles de los experimentadores	Evento objetivo	Evento subjetivo
Sitio	Evento global	Evento local
Protocolo experimental	Evento real	Evento artificial
Operacionalizaciones de los constructos	Límites de operacionalizaciones conocidas	Evento limitado a la forma de aplicar los tratamientos y medir las respuestas
Propiedades de las poblaciones	Límites de poblaciones conocidas	Evento limitado a ciertas propiedades de las poblaciones

Tabla 6.2: Conocimiento ganado en la replicación al variar los elementos de la situación experimental

Por ejemplo, si se identifican distintos cambios en el protocolo experimental y los resultados son reproducibles, el grado de verificación que se obtiene es que los resultados son reales y no artificiales (producto de la configuración del protocolo experimental). Si por el contrario, se identifican distintos cambios en los elementos de la situación experimental de las replications y los resultados no son reproducibles es difícil, sino imposible, identificar el o los elementos variados que pueden ser causa de las diferencias entre los resultados.

La semejanza entre replications ofrece una idea sobre la cantidad de cambios que hay en las situaciones experimentales de las replications de una partición. No obstante cuando se tienen múltiples replications con distintos cambios es difícil, sino imposible, identificar el nivel de semejanza que hay entre las replications, esto debido a que la información de las situaciones experimentales aumenta de acuerdo al número de replications que se tengan en una partición lo que hace intratable conocer a simple vista las semejanzas entre replications. De nuevo es necesario una ayuda que facilite al investigador conocer la semejanza entre replications según los cambios identificados en las situaciones experimentales.

Una manera de conocer el nivel de semejanza entre replications es usar algún tipo de distancia que represente la similitud entre replications de una partición, donde las replications con mayor número de cambios estén a una distancia mayor que aquellas con un número de cambios menor. Las distancias resultantes pueden usarse para representar de manera gráfica la similitud que hay entre las distintas replications de una partición respecto a sus situaciones experimentales.

En la tabla 6.3 se muestra un ejemplo donde distintos cambios relacionados con el protocolo experimental de tres replicaciones (los mostrados en la tabla 6.1) se han transformado en una matriz de valores numéricos que representan las similitudes entre estas replicaciones.

	r1	r2	r3
r1	0		
r2	0.3333	0	
r3	0.0333	0.3	0

Tabla 6.3: Ejemplo de similitudes representadas como distancias entre tres replicaciones respecto al protocolo experimental

Distancias con valores cercanos a 1 indican una mayor dis-similitud mientras que distancias con valores próximos a 0 indican mayor similitud. En el caso de la replicación *r2* es menos similar a las replicaciones *r1* y *r3*, mientras que las replicaciones *r1* y *r3* con menos cambios se asemejan más entre sí. Al observar los cambios registrados referentes al protocolo experimental (ver tabla 6.1), las replicaciones *r1* y *r3* difieren sólo en el elemento denominado *tratamientos por sesión* donde en *r1* se aplicaron tres técnicas a un programa, mientras que en *r3* se aplicaron dos técnicas a un programa, el resto de elementos se mantiene sin cambios. Por otra parte, la replicación *r2* contiene más cambios respecto a las otras dos replicaciones. En este ejemplo, respecto a la replicación *r1*, en *r2* se han variado los elementos: *técnica de análisis de datos*, *diseño experimental*, *orden de aplicación de los tratamientos*, *tratamientos por sesión* y *tareas por sujeto por sesión*. De manera similar la replicación *r2* difiere de *r3*.

Por cada sección de la situación experimental se transforman los cambios registrados en valores numéricos hasta obtener una sola matriz similar a la de la tabla 6.3 donde se representen las situaciones experimentales. Usando esta matriz, se puede obtener una representación gráfica. La figura 6.4 muestra un ejemplo de dos representaciones gráficas de las similitudes entre las replicaciones *r1*, *r2* y *r3* respecto a las situaciones experimentales. En (6.4a) se utiliza un dendrograma donde se van agrupando objetos similares, en este caso, replicaciones. En esta figura las replicaciones *r2* y *r3* que son más similares entre sí forman un grupo que se une a una distancia de casi 0.4 unidades, mientras que la replicación *r1* se une a este grupo a una distancia de casi 0.7 que indica una mayor dis-similitud en *r1* respecto a las replicaciones *r2* y *r3*. Es decir este valor próximo a 0.7 indica un mayor número de cambios en los elementos de la situación experimental en la replicación *r1*.

Por otra parte, en la figura 6.4b se usa un plano de dos dimensiones para representar las similitudes de las replicaciones. Los objetos más cercanos entre sí, en este caso replicaciones, indican una mayor similitud, mientras que los más lejanos entre sí indican mayor dis-similitud. En esta figura, las replicaciones *r2* y *r3* se encuentran más cercanas entre sí lo que indica que aunque ambas replicaciones contienen cambios en la situación experimental éstas se asemejan más entre sí, a diferencia de la replicación *r1* ubicada a una distancia mucho mayor². Ambas representaciones gráficas facilitan la identificación del nivel de similitud entre situaciones experimentales de las replicaciones pertenecientes a una partición.

²En el apéndice B se describe detalladamente distintos tipos de medidas de proximidad así como técnicas para representar las similitudes entre varias replicaciones.

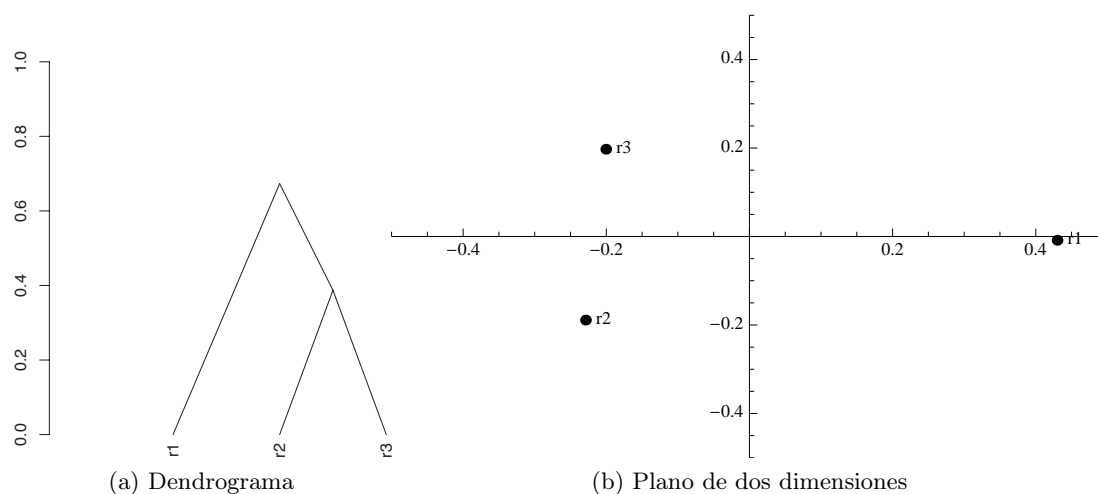


Fig. 6.4: Dos tipos de gráfico que representan las similitudes entre tres replicasiones

6.3 Variabilidad de los Tratamientos

Cuando se tienen varias replicasiones es habitual intentar combinarlas usando alguna técnica de agregación de resultados. De manera general, las técnicas de agregación se usan para estimar un tamaño de efecto global a partir del tamaño de efecto³ estimado individualmente en cada estudio o experimento (en este caso replicación), y para conocer si los resultados de varios experimentos son homogéneos i.e. si son semejantes entre sí, o heterogéneos i.e. si se observan diferencias entre los resultados.

Se dice que los resultados son homogéneos si la distribución de tamaños de efecto estimados se asemejan entre sí, lo que se interpreta como que hay un tamaño de efecto común entre el conjunto de tamaños de efecto de las replicasiones estudiadas. Por otra parte, se dice que los resultados son heterogéneos cuando en la distribución de tamaños de efecto estudiados hay uno o varios tamaños de efecto que no se asemejan al resto, lo que se interpreta como que no hay un efecto común en el conjunto de tamaños de efecto estudiados. En la literatura se encuentran distintas técnicas que permiten evaluar la heterogeneidad de un conjunto de tamaños de efecto. En el apéndice B se describen algunas técnicas usadas para evaluar la heterogeneidad en un conjunto de experimentos, en este caso, replicasiones.

Los resultados de un meta-análisis suelen representarse de manera gráfica. En la figura 6.5 se muestra una técnica de agregación típica⁴ con la información de nueve replicasiones que examinan la efectividad de dos técnica así como un gráfico con los tamaños de efecto estimados y el tamaño de efecto global estimado.

En esta figura el tamaño de efecto estimado en cada replicación se representa por un cuadrado, las líneas horizontales indican el intervalo donde se cree puede estar el tamaño de efecto verdadero. El tamaño de los cuadrados indican la precisión del efecto estimado, cuanto más grande sea el cuadrado se tiene una mayor precisión en el efecto estimado y el intervalo donde se cree puede estar el tamaño de efecto verdadero será estrecho. Por el contrario, cuanto más pequeño sea el cuadrado, se tiene menor precisión sobre el efecto

³Como se mencionó en el capítulo 4, un tamaño de efecto es representado por un coeficiente que mide las diferencias entre dos tratamientos.

⁴En el apéndice B se describe a detalle la técnica de agregación representada en esta figura.

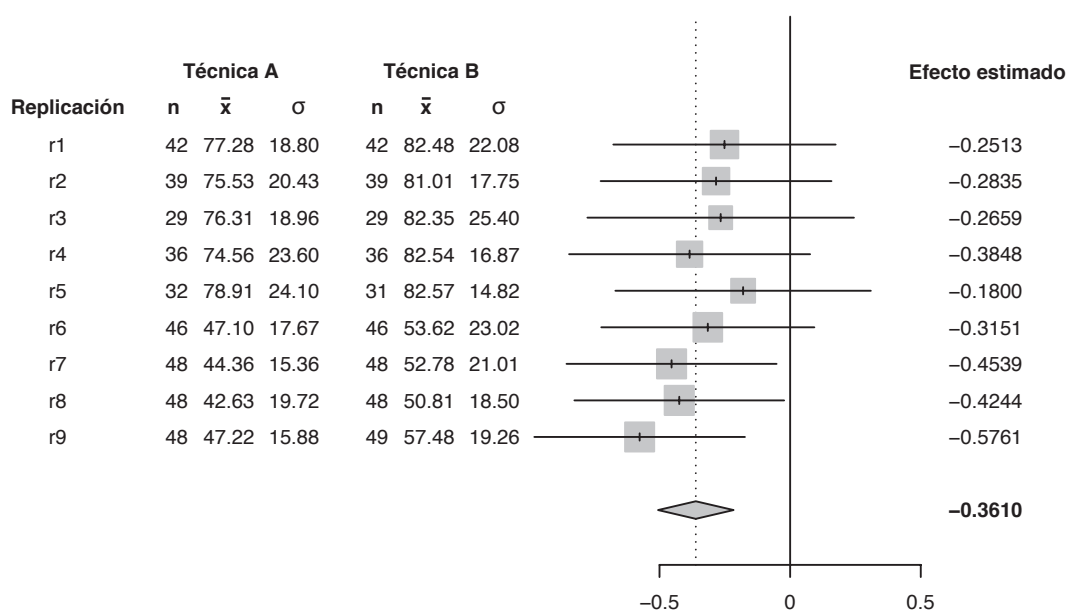


Fig. 6.5: Ejemplo de agregación de resultados de nueve replicaciones

estimado y el intervalo donde se cree puede estar el tamaño de efecto verdadero será ancho. El rombo que se encuentra en la parte inferior indica el efecto global estimado y su anchura al igual que las líneas horizontales indica el intervalo donde se cree está el efecto global verdadero.

De manera general, de acuerdo con la información de la figura 6.5 se observa que los tamaños de efecto estimados en las replicaciones del ejemplo se asemejan entre sí, obteniendo como resultado que la técnica B es más efectiva que la técnica A. Por último el efecto global estimado indica que en estas replicaciones la técnica B es más efectiva.

Las técnicas de agregación son útiles para estimar tanto tamaños de efecto individuales como para estimar un tamaño de efecto global. No obstante, estas técnicas son ciegas a los valores observados de la variable de respuesta para cada tratamiento. Por ejemplo, en dos replicaciones $r1$ y $r2$ donde se evalúa la efectividad de dos técnicas A y B, se ha obtenido un tamaño de efecto de -0.3848 para $r1$ y un tamaño de efecto -0.4539 para $r2$. Ambos tamaños de efecto indican que la técnica B es más efectiva que la técnica A.

Las variables de respuesta obtenidas en $r1$ fueron una efectividad del 74.56% para la técnica A y una efectividad del 82.54% para la técnica B. Por otra parte en la replicación $r2$ se obtuvo una efectividad del 44.36% en la técnica A y una efectividad del 52.78% para la técnica B. Aunque los tamaños de efecto indiquen una homogeneidad ya que se asemejan entre sí ($r1$ -0.38 y $r2$ -0.45), las variables de respuesta en ambas replicaciones varían de manera notoria. En $r2$ las variables de respuesta i.e. las magnitudes de los tratamientos (técnica A y B) son menores respecto a $r1$. En la figura 6.6 se muestra esta situación. En la figura 6.6a se muestra la distribución de tamaños de efecto de estas dos replicaciones, los cuales se asemejan mucho entre sí i.e. ambos efectos indican que la técnica B es más efectiva que la técnica A.

Por otra parte en las figuras 6.6b y 6.6c se muestran dos gráficas con las variables de respuesta por cada técnica. En estas dos gráficas se observa que algo afectó a las variables de respuesta de la replicación $r2$ porque las magnitudes en ambas técnicas son

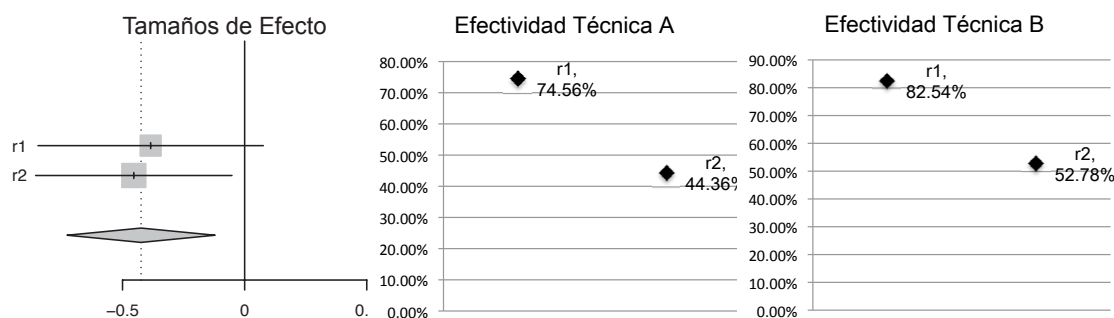


Fig. 6.6: Tamaños de efecto de dos replicaciones y valores de las variables de respuesta de los tratamientos usados en $r1$ y $r2$

mucho menores que las magnitudes obtenidas en $r1$. Para apreciar este tipo de diferencias de resultados entre replicaciones se hace necesario contar con otro tipo de herramienta.

No basta con usar técnicas de agregación para verificar resultados experimentales, se deben realizar otros tipos de análisis no sólo a nivel de tamaños de efecto como en las técnicas de agregación, sino también a nivel de cada tratamiento. ¿Qué sucede cuando los resultados de la agregación indican homogeneidad pero hay variaciones en las magnitudes de los tratamientos?

En la figura 6.5 se muestra otro ejemplo con un número mayor de replicaciones. Como se observa en la figura, las replicaciones $r1$ a $r5$ y $r6$ a $r9$ difieren de manera importante en las magnitudes de los tratamientos. No obstante esta situación no se observa en el gráfico de la figura 6.5, donde la distribución de efectos estimados se asemejan entre sí.

Si además de examinar los tamaños de efecto estimados se examina individualmente las magnitudes de cada tratamiento se generará más conocimiento. Si se usa un gráfico de dos dimensiones para representar de manera sucesiva los valores de la técnica A de las nueve replicaciones se aprecia fácilmente el cambio de magnitud, tal como se muestra en la figura 6.7.

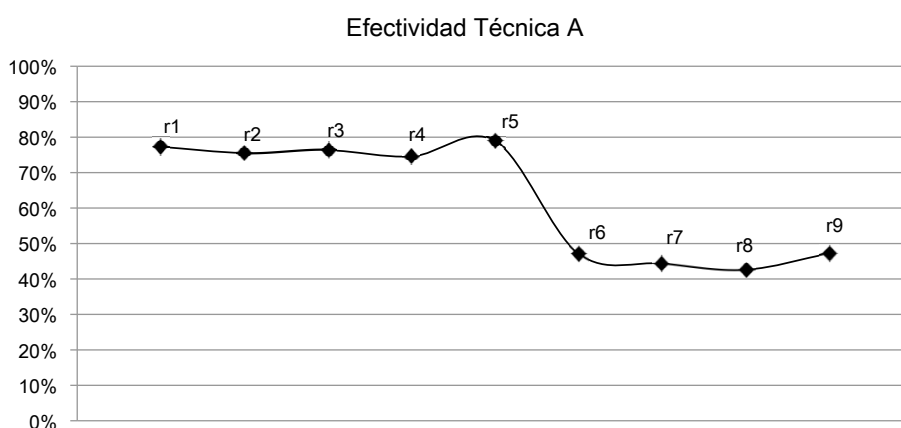


Fig. 6.7: Gráfico de dos dimensiones que representa la efectividad de la técnica A aplicada en nueve replicaciones

Este tipo de gráfico es útil para examinar y conocer la variación de un tratamiento entre replicaciones. Por ejemplo si se tienen varias replicaciones literales este gráfico

facilita examinar y conocer la variabilidad natural de la variable de respuesta de un tratamiento.

Para decidir si la variabilidad en un tratamiento es natural o producto de algún cambio en los elementos de la situación experimental es necesario identificar los límites de la variabilidad natural de la variable de respuesta para el tratamiento.

Siguiendo con este mismo tipo de gráfico, en la figura 6.8 se muestra otro ejemplo donde se examina la variabilidad de la técnica A en seis replicaciones. En este ejemplo se han especificado límites que indican si los valores de la técnica corresponden a una variabilidad natural o si corresponden a la modificación de algún elemento de la situación experimental. En este caso los límites permiten estimar una variabilidad natural promedio en la efectividad de la técnica A en torno al 42%. Estos límites también permiten conocer valores que no se encuentren dentro de la variabilidad natural especificada, como sucede con la replicación r_4 donde el valor tan alto arrojado en la efectividad de la técnica A de esta replicación pudo deberse a uno o varios cambios hechos en la situación experimental.

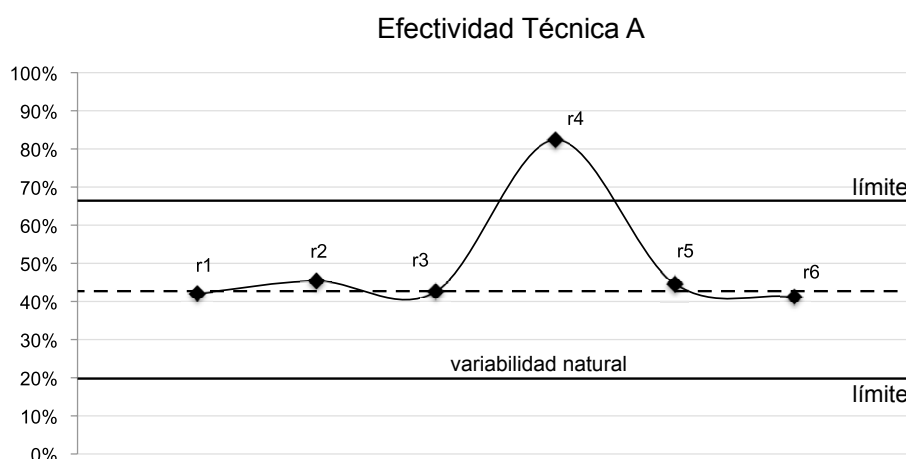


Fig. 6.8: Ejemplo de variabilidad de la técnica A en seis replicaciones

Gráficos similares al aquí descrito son usados rutinariamente en empresas del sector industrial. En este tipo de gráficos también conocidos como *gráficos de control* se representan valores de algún tipo de medición realizada durante el funcionamiento de un proceso continuo y que además sirven para controlar dicho proceso. Bajo condiciones normales del proceso, los valores fluctúan dentro de un intervalo determinado, en caso de haber alguna anomalía en el proceso los valores salen de este intervalo.

Para construir un gráfico de control es necesario estimar el promedio así como los límites bajo los cuales las observaciones se consideran como fluctuaciones comunes del proceso. Se asume que las observaciones se distribuyen de acuerdo a una distribución normal, por lo que se dirá que el proceso está bajo control si las observaciones obtenidas siguen las características de una distribución normal. Aproximadamente el 67% de las observaciones han de encontrarse dentro de la primera desviación ($\pm 1\sigma$), aproximadamente el 95% de las observaciones han de encontrarse dentro de la segunda desviación ($\pm 2\sigma$) y aproximadamente el 99% de las observaciones han de encontrarse dentro de la tercera desviación ($\pm 3\sigma$). En esta última desviación se delimita el proceso con los límites superior UCL (*Upper Control Limit*, Límite de Control Superior) e inferior LCL (*Lower*

Control Limit, Límite de Control Inferior)⁵.

Para examinar la variabilidad de los tratamientos entre replicaciones se propone usar gráficos de control individual. Existen diferentes tipos de gráficos de control, el más conocido y usado en empresas del sector industrial consiste en obtener muestras en un periodo determinado y de una tamaño determinado de observaciones de alguna variable que se esté midiendo en el proceso. A cada muestra se calcula su promedio y se representa por una observación en el gráfico de control. Cuando no se dispone de numerosas muestras, se utiliza otro tipo de gráfico de control conocido como gráfico de control individual. En este gráfico cada observación representa un valor individual.

El gráfico de control individual se propone usarlo en ISE para examinar la variabilidad de los tratamientos aplicados en las replicaciones, esto debido a que en ISE no se dispone de un gran número de observaciones como en el caso de las empresas del sector industrial, y no se dispone de un número fijo de observaciones para constituir las muestras. En ISE cada replicación suele tener tamaños de muestra distintos.

Existen diferentes conjuntos de reglas que se usan para evaluar la tendencia de las observaciones en los gráficos de control y determinar si el proceso se encuentra fuera de control, un ejemplo de este tipo de reglas se muestra a continuación:

1. Cualquier observación fuera de los límites de la tercera desviación.
2. Dos de tres observaciones sucesivas fuera de la segunda desviación.
3. Cuatro de cinco observaciones sucesivas se encuentran fuera de la primera desviación.
4. Ocho observaciones consecutivas que se encuentren dentro de las primeras dos desviaciones pero alejadas de la línea central.

Respecto a los distintos conjuntos de reglas que hay para evaluar la tendencia de las observaciones en los gráficos de control, se tiene común acuerdo en la primera regla, que dice, cualquier observación fuera de los límites de la tercera desviación indica que el proceso está fuera de control. Ya que hay común acuerdo en esta regla, **se propone usar esta regla de las tres desviaciones como punto de partida para evaluar la variabilidad de los tratamientos en un conjunto de replicaciones y determinar si los resultados son o no reproducibles.**

El uso de gráficos de control en ISE acompañado de alguna técnica de agregación de resultados ofrece al investigador mayor información respecto a si los resultados en las replicaciones son o no verificables y por tanto reproducibles. En el capítulo 9 así como en el apéndice B se describe con detalle como emplear este tipo de gráficos para examinar la variabilidad de los tratamientos entre replicaciones.

6.4 Grado de Verificación Logrado

El grado de verificación logrado en las replicaciones de una partición depende de dos condiciones, 1) de los cambios identificados en las situaciones experimentales, y 2) de los resultados obtenidos tras examinar la variabilidad de los tratamientos con los gráficos de control y tras examinar con una técnica de agregación la heterogeneidad i.e. la distribución de tamaños de efecto.

Si en los gráficos de control, la variabilidad de los tratamientos se encuentra dentro de los límites de control (tres desviaciones), y la distribución de los

⁵En el apéndice B se describe detalladamente el funcionamiento de un gráfico de control que puede usarse en ISE para examinar la variabilidad de tratamientos entre replicaciones.

tamaños de efecto se considera homogénea (semejantes entre sí), se propone denominar consistentes los resultados en las replicaciones de la partición ya que han sido reproducidos.

Así pues, el grado de verificación logrado dependerá de dos criterios, 1) si hay o no cambios en la situación experimental de las replicaciones de una partición, y 2) si los resultados de la partición son o no consistentes. De acuerdo a estos criterios, en una partición se presentan cuatro posibles situaciones:

1. Situación experimental sin cambios, con resultados consistentes
2. Situación experimental con cambios, con resultados consistentes
3. Situación experimental sin cambios, con resultados inconsistentes
4. Situación experimental con cambios, con resultados inconsistentes

La primera situación se presenta cuando se tienen replicaciones literales (se mantienen sin cambios los elementos de la situación experimental) que proporcionan todas ellas resultados consistentes. El grado de verificación que se logra es que los resultados observados no son fortuitos producto de un error de tipo I dado que se ha observado n veces que el resultado no se presenta por coincidencia. En esta situación también se conoce la variabilidad natural de los tratamientos aplicados en las replicaciones de la partición.

En la segunda situación las replicaciones de la partición pueden tener distintos niveles de similitud i.e. puede haber distintos tipos de cambios en los elementos de la situación experimental. Al observar resultados consistentes en esta situación, el grado de verificación logrado dependerá de los cambios hechos en los diferentes elementos de la situación experimental. Por ejemplo, si en las replicaciones se han variado los roles de los experimentadores, el sitio y el protocolo experimental, se verifica que los resultados son objetivos (independientes de los experimentadores), globales (independientes del sitio) y reales (independientes del protocolo experimental empleado).

La tercera situación se presenta cuando se tienen replicaciones literales (mismas situaciones experimentales) y se observan resultados inconsistentes, ya sea debido a la variación no natural en alguno o ambos tratamientos o a que la distribución de los tamaños de efecto no es homogénea sino heterogénea. En esta situación algún elemento de la situación experimental pudo influir en los resultados, sin embargo al no identificar cambios en la situación experimental es difícil tratar de identificar posibles causantes de estos resultados. Una interpretación puede ser que no se haya hecho apropiadamente el registro de cambios en los elementos de la situación experimental. Otra posible interpretación es que los resultados son así de inconsistentes. En esta situación se recomienda revisar si hubo cambios en algún elemento de la situación experimental y éste no se registró. Puede presentarse la situación de encontrar a posteriori un elemento que se cree pudo influir en los resultados. Por ejemplo, tras realizar varias replicaciones literales y obtener resultados inconsistentes, se encuentra que la motivación, variable no identificada hasta ahora, puede influir en los resultados de alguna de estas replicaciones.

La cuarta situación se presenta cuando se identifican cambios en los elementos de la situación experimental y se observan resultados inconsistentes en las replicaciones de una partición. En esta situación probablemente alguno de los cambios identificados en la situación experimental produce el cambio en los resultados.

En la tabla 6.4 se muestran estas cuatro situaciones con el tipo de aprendizaje que se obtiene, recomendaciones a seguir y el tipo de verificación que se logra en caso de obtener resultados consistentes en una partición.

A partir de los grados de verificación logrados en cualquiera de las primeras dos

Cambios en situación experimental	Tipo de resultados	Aprendizaje	Recomendación	Verificación
No	Consistentes	Variabilidad Natural	Avanzar en la verificación de resultados con otros tipos de replicación	Resultados no fortuitos
Sí	Consistentes	Conocimiento respecto a los cambios de la configuración experimental	Dependiendo del tipo de replicación realizado, avanzar de forma sistemática en la verificación de resultados usando otro tipos de replicación	Dependiendo del cambio: -si experimentadores: resultado objetivo -si sitio: resultado global -si protocolo: resultado real -si operacionalizaciones: límites en aplicación de tratamientos y medición de respuestas -si poblaciones: limitado a ciertas poblaciones
No	Inconsistentes	Conocimiento de condiciones experimentales que influyen en resultados	Buscar cambios en situación experimental	-
Sí	Inconsistentes	Conocimiento de condiciones experimentales que influyen en resultados	Buscar e identificar posibles condiciones causantes de resultados inconsistentes	-

Tabla 6.4: Situaciones presentadas en una partición, aprendizaje y tipo de verificación que se obtiene

situaciones es posible obtener algún tipo de conocimiento. En la primera situación, como se mencionó, se gana el conocimiento sobre la variabilidad natural de los resultados de las replicaciones en una partición. No obstante, se deben llevar a cabo otros tipos de replicación, con el fin de verificar que los resultados observados no son artificiales.

En la segunda situación, como se mencionó, se obtienen distintos tipos de conocimiento, dependiendo de los cambios que se realicen en la situación experimental, se conocen los elementos de la situación experimental (por tanto las variables asociadas) que pueden variarse sin que influyan en los resultados experimentales. También se conoce el tipo de verificación logrado. Por ejemplo, si se varían los experimentadores se obtiene como conocimiento que el resultado es objetivo independiente de los experimentadores, si varía el sitio, se obtiene como conocimiento que los resultados son globales independientes del sitio, si se varía el protocolo experimental se obtiene como conocimiento que los resultados obedecen a una realidad, si se varían las operacionalizaciones se obtiene como conocimiento los límites en que los resultados se generalizan a otras formas de operacionalizar los constructos, si se varían las poblaciones se obtienen como conocimiento los límites en que los resultados se generalizan a otros tipos de poblaciones.

Como se observa en la tabla 6.4, a partir de la obtención de resultados consistentes, es posible obtener algún tipo de conocimiento. Por ejemplo, si se tienen resultados consistentes en seis replicaciones de una partición donde se evalúa la efectividad de dos técnicas A y B, es posible conocer los resultados obtenidos, el tipo de verificación logrado, el nivel de variaciones hechas en la situación experimental así como los tipos de replicaciones que forman parte de la partición.

Asumiendo que en estas replicaciones se hayan variado todos los elementos de la situación experimental referentes a experimentadores, sitio, operacionalizaciones así como las poblaciones, y se hayan variado parcialmente algunos elementos del protocolo experimen-

tal, se obtiene como conocimiento que los resultados son objetivos, globales, semi-reales, con conocimiento en la forma de aplicar los tratamientos y medir los constructos, y con conocimiento de las poblaciones equivalentes. Al variar en estas replicaciones los distintos elementos de la situación experimental se realizan los distintos tipos de replicaciones (nativa, conjunta, ajena, interna, externa, artificial, operacional y poblacional) que se identificaron con la tipología propuesta (capítulo 5).

Todo este conocimiento puede ser empaquetado y organizado. Siguiendo con el ejemplo, en la tabla 6.5 se muestra el conocimiento obtenido de seis replicaciones que conforman una partición donde sus resultados son consistentes y se han variado distintos elementos de la situación experimental.

Información de la Partición			
Partición:	Efectividad/ Métrica A/ Técnica A vs Técnica B		
Experimentos de la partición:	6 (r1, r2, r3, r4, r5, r6)		
Situación experimental:	Con Cambios en varios elementos		
Resultados observados:	Consistentes		
Información de la comparación			
Comparación:	La efectividad de la técnica A es mejor respecto a la técnica B.		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	100%	Objetivo	Nativa, Conjunta, Ajena
Sitio:	100%	Global	Interna, Externa
Protocolo experimental:	77%	Semi-real	Artificial
Operacionalizaciones de los constructos:	100%	Con conocimiento en la forma de aplicar los tratamientos y medir los resultados	Operacional
Poblaciones:	100%	Con conocimiento de poblaciones equivalentes	Poblacional

Tabla 6.5: Ejemplo de conocimiento obtenido respecto al grado de verificación de seis replicaciones

Por el contrario, cuando se presentan resultados inconsistentes en una partición, es posible identificar posibles condiciones o elementos de la situación experimental que se cree influyen en los resultados. En esta situación también es posible obtener conocimiento ya que se identifican posibles variables que sirven como punto de partida para generar nuevas hipótesis de trabajo que pueden ser testadas en replicaciones posteriores.

6.5 Identificación de Posibles Causas de Resultados no Reproducibles

Cuando los resultados en alguna de las particiones son inconsistentes y por tanto no se han reproducido, se propone llevar a cabo un análisis con el fin de facilitar la identificación de posibles elementos de la situación experimental que pueden influir en los resultados.

A través del registro de cambios en las situaciones experimentales de varias replicaciones (discutido en la sección 6.2) es posible identificar el o los elementos que se creen influyeron en los resultados de una partición. No obstante, al haber tantos elementos en la situación experimental, es imposible llevar un control mental de los elementos que se cree afectan a los resultados. Se requiere algún tipo de ayuda que facilite al investigador tanto registrar posibles causantes así como llevar a cabo un análisis de las causas de resultados no reproducibles en una partición.

Existen técnicas que facilitan el análisis de problemas y sus soluciones afines a la calidad de procesos, productos o servicios. **Se propone utilizar diagramas de causa y efecto con el fin de facilitar el análisis e identificación de los elementos de la situación experimental que parecen influir en los resultados de las replicasiones de una partición.** Los diagramas de causa y efecto facilitan el análisis y registro de causas que se cree ocasionan algún tipo de problema. En la figura 6.9 se muestra un ejemplo donde vemos adaptada esta técnica a la ISE.

En la parte izquierda del diagrama de esta figura se encuentran las posibles causas, mientras que en la parte derecha el efecto observado. El efecto puede representarse con los resultados de los gráficos de control o de la técnica de agregación. En este diagrama, las causas principales están representadas por los distintos elementos que conforman la situación experimental. En este diagrama se pueden escribir anotaciones para identificar posibles condiciones o elementos que se cree ocasionan resultados inconsistentes. Los elementos identificados con este tipo de técnica sirven para generar hipótesis que pueden ser testadas en replicasiones posteriores. En el capítulo 9 así como en el apéndice B se describe con detalle como emplear este tipo de herramienta para identificar y analizar las condiciones o elementos candidatos a ocasionar resultados no reproducibles en una partición.

6.6 Comparando Replicasiones

Una vez presentadas las ideas propuestas, a continuación se entrelazan estas ideas de modo que hagan posible llevar a cabo comparaciones entre replicasiones.

Recordando la problemática que se desea resolver. Actualmente los intentos de combinar replicasiones en ISE para extraer algún tipo de conocimiento no han sido fructíferos. En los casos donde es posible combinar algún conjunto de replicasiones porque sus resultados se asemejan entre sí, no se discute el grado de verificación alcanzado en esas replicasiones, simplemente se agregan los resultados. ¿Son los resultados observados independientes de los investigadores o del sitio donde se realizaron las replicasiones? ¿Son los resultados reales o artificiales producto de la configuración del protocolo experimental empleado?. Por otra parte, en los casos donde no es posible combinar un conjunto de replicasiones debido a las diferencias de resultados, no se es capaz de extraer algún tipo de conocimiento (por ejemplo, que los resultados son artificiales) e ingenuamente se concluye que debido a las diferencias de resultados y a los cambios hechos en las replicasiones, éstas no son comparables. Resultando, por tanto, las replicasiones inútiles pues no contribuyen a obtener ningún tipo de conocimiento.

En conjunto, las ideas presentadas en este capítulo aportan una solución a la problemática enunciada. En la figura 6.10 se muestran las ideas propuestas interconectadas que unidas permiten comparar replicasiones y obtener conocimiento de ellas.

Como primer paso para comparar replicasiones satisfactoriamente, es necesario identificar las unidades comparables comunes entre las replicasiones que se desean comparar. Una unidad comparable se conforma por un constructo y operacionalización de efecto y un par de constructos y operacionalizaciones de causa de una replicación. Las unidades comunes entre sí se agrupan en conjuntos denominados particiones.

Sin este concepto de unidad comparable se intentaría comparar replicasiones de manera completa lo que es inviable ya que seguramente se deseche replicasiones porque, por ejemplo, no compartan el cien por ciento de tratamientos cuando en realidad poseen

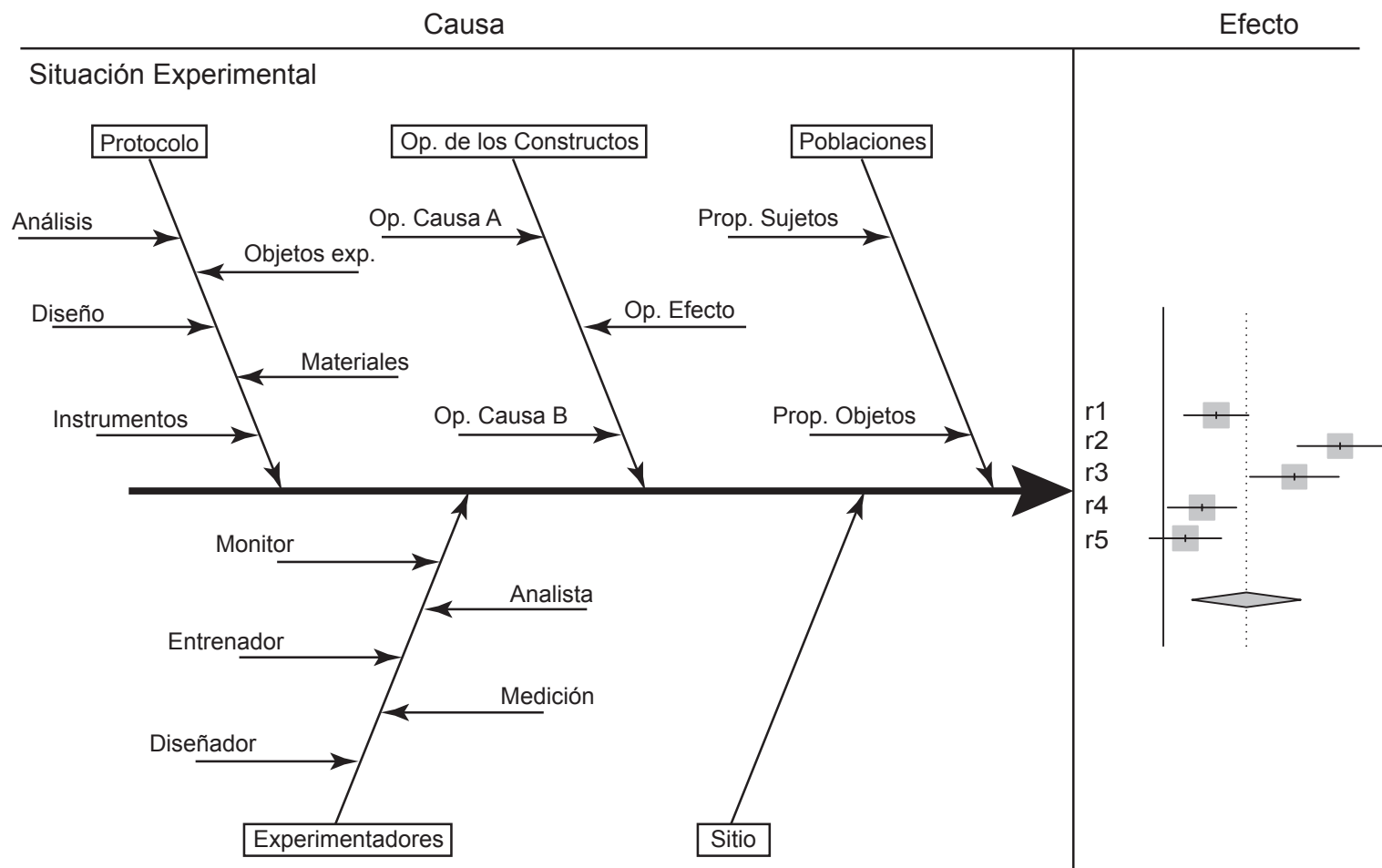


Fig. 6.9: Ejemplo de adaptación de técnica para analizar causas de resultados inconsistentes

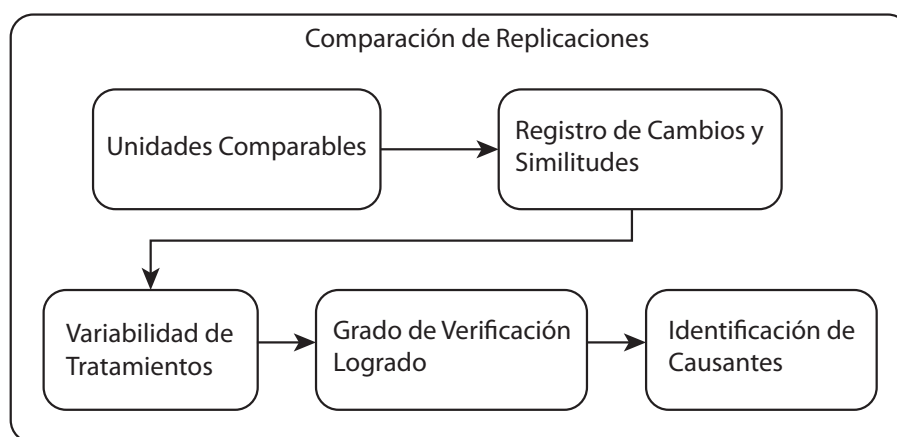


Fig. 6.10: Secuencia de ideas propuestas que en conjunto hacen posible la comparación de replications

unidades comparables comunes que pueden agruparse en particiones diferenciadas de modo que se realicen comparaciones intra-partición.

Una vez que se dispone de alguna partición el *modus operandi* tradicional de la ISE llevaría a aplicar técnicas de agregación de resultados para comparar el tamaño de efecto global. Así no es posible conocer el grado de verificación logrado por estos resultados y es difícil identificar posibles condiciones que ocasionan resultados no reproducibles. Para conocer el grado de verificación alcanzado en una partición, es necesario identificar los cambios hechos en las situaciones experimentales.

La segunda idea propone registrar exhaustivamente los cambios hechos en las situaciones experimentales de las replications de una partición. Así en caso de obtener resultados reproducibles se podrá conocer su grado de verificación o en caso contrario identificar posibles causantes de resultados no reproducibles en la partición. Tras el registro de cambios se propone calcular la semejanza que hay entre las situaciones experimentales de las replications en una partición. Las semejanzas ofrecen una idea sobre la cantidad de cambios que hay en las situaciones experimentales de una partición.

Una vez identificados los cambios en las situaciones experimentales de una partición, el siguiente paso consiste en examinar los resultados experimentales de la partición. Es aquí donde interviene la tercera idea propuesta. No basta con emplear técnicas de agregación de resultados para determinar si los resultados se han o no reproducido porque como se describió anteriormente, estas técnicas son ciegas a los valores observados de la variable de respuesta de cada tratamiento. Se debe examinar también la variabilidad de las variables de respuesta de los tratamientos aplicados en las replications de la partición. Se propone además del uso de técnicas de agregación, utilizar gráficos de control para examinar la variabilidad de las variables de respuesta de los tratamientos.

Ya identificados los cambios en las situaciones experimentales y examinados los resultados experimentales de la partición (mediante gráficos de control y técnicas de agregación), la cuarta idea propuesta ayuda al investigador a identificar el tipo de conocimiento que se puede obtener dependiendo de si hubo o no cambios en las situaciones experimentales y de si los resultados de la partición son o no consistentes. En caso de obtener resultados consistentes nuestra propuesta facilita la identificación del grado de verificación logrado. Por el contrario si se observaron resultados inconsistentes es posible

identificar causantes que se cree influyen en los resultados de una partición.

La última idea propuesta complementa a la anterior y su función es ayudar al investigador a identificar condiciones o elementos de la situación experimental candidatos a ocasionar resultados no reproducibles en una partición. Esta propuesta permite obtener conocimiento sobre posibles condiciones que sirven para generar nuevas hipótesis y testarlas en replications futuras.

Como se observa, cada propuesta satisface un propósito particular y al combinarlas aportan una solución global que facilita la comparación de replications. Utilizando la ideas de este marco, en el siguiente capítulo se describe un procedimiento que puede usarse para llevar a cabo comparaciones entre replications.

Capítulo 7

Procedimiento para Comparar Replicaciones en ISE

RESUMEN: Usando como referencia el marco para la comparación de replicaciones, en este capítulo se describe un procedimiento que permite al investigador comparar conjuntos de replicaciones para obtener algún conocimiento de ellas.

Una vez desarrolladas las primeras tres tareas de la solución, que son: la clarificación de conceptos de replicación en ISE, desarrollo de una tipología de replicaciones para la ISE y desarrollo de un marco conceptual para comparar replicaciones en ISE, en este capítulo se aborda la cuarta tarea, desarrollar un procedimiento que ayude al investigador en la explotación y comparación de conjuntos de replicaciones con el fin de obtener conocimiento que sea de utilidad tanto por el profesional como para el investigador.

Este procedimiento se basa en las ideas clave del marco conceptual presentado en el capítulo anterior. La finalidad de este procedimiento es ayudar al investigador a sintetizar la información de un conjunto de replicaciones para obtener conocimiento de ellas. El conocimiento obtenido es entonces empaquetado en piezas que sean útiles tanto para el profesional como para el investigador. Las piezas de conocimiento contienen información sobre las unidades comparables que hay en común entre varias replicaciones.

Para el profesional, el conocimiento obtenido es de utilidad porque a partir de éste se toman decisiones mucho más objetivas sobre algún aspecto de la construcción del software o sobre la adopción de alguna nueva tecnología software. Con este conocimiento se evita lo que a menudo sucede en nuestra disciplina, adoptar tecnologías nuevas sin contar con evidencia convincente que éstas serán efectivas e ignorar aquellas que han demostrado ser efectivas a pesar de la publicación de datos que así lo demuestran (Zelkowitz et al., 2003).

Para el investigador, el conocimiento obtenido es de utilidad porque permite conocer los distintos cambios realizados en las replicaciones comparadas, las condiciones donde se han reproducido los resultados o los límites donde se ha observado algún efecto. Este conocimiento sirve también como referente para determinar qué otros tipos de replicación pueden realizarse con el fin de conocer y estudiar nuevas condiciones que influyan en el proceso de construcción del software.

Este procedimiento se compone de varias partes donde cada parte cumple una función particular. Para representar las distintas partes de este procedimiento se ha usado como referencia una extensión del lenguaje UML (*Unified Modeling Language*, Lenguaje de Modelado Unificado) desarrollada por Eriksson and Penker (Eriksson y Penker, 1998). Esta

extensión suele usarse para modelar procesos de negocio, no obstante, también permite modelar procesos genéricos. En este sentido, se ha elegido esta extensión para modelar este procedimiento y sus partes como un proceso conformado por sub-procesos.

De acuerdo a esta extensión, un diagrama de proceso es un diagrama de actividades UML con un conjunto de estereotipos que describe tanto las actividades realizadas dentro de los procesos como su interacción. Un proceso toma uno o más tipos de entrada que se utilizan para transformar y producir uno o varios tipos de salidas, este elemento representa una actividad UML estereotipada como proceso. A su vez, un proceso puede contener uno o más sub-procesos que contienen como unidad atómica una o más actividades. Las diferentes entradas y salidas de un proceso reciben el nombre de recursos y se representan por instancias de clases UML. En esta extensión existen cuatro tipos de recursos:

- Físico. Representa una entidad material real que ocupa un volumen de espacio
- Abstracto. Representa una idea o concepto
- Información. Representa hechos o conocimientos acerca de algún tipo de objeto
- Personas. Representa un ser humano que interactúa con el proceso

En la figura 7.1 se muestra una vista general de este procedimiento modelado como un proceso. Este procedimiento usa sólo dos tipos de recursos que son de tipo *información* y de tipo *personas*. Como entrada, se recibe la información de un conjunto de replications. La información de las replications puede estar disponible en diferentes tipos de formato como son: artículos, manuscritos, informes técnicos, paquetes de replicación, o datos experimentales de las replications almacenados en repositorios. Este procedimiento es controlado por uno o varios investigadores. Como suministro, se utiliza la tipología de replications, el marco conceptual así como varias técnicas y herramientas descritas en el apéndice B. Como salida, se obtiene conocimiento que es empaquetado en una o varias piezas.

Este procedimiento se compone de dos sub-procesos, *generador de particiones* y *comparador de replications*. Para obtener conocimiento de un conjunto de replications, primero es necesario identificar las unidades comparables comunes entre las replications a comparar. El *generador de particiones* recibe como entrada tres o más replications¹ para agruparlas en particiones de acuerdo a las unidades comparables que hay en común en las replications a comparar. Como se discute en el capítulo 6 una partición se conforma por un constructo y operacionalización de efecto y un par de constructos y operacionalizaciones de causa.

El segundo sub-proceso (*comparador de replications*) recibe como entrada una partición y la transforma en una o varias piezas de conocimiento. Este sub-proceso examina los cambios de las replications que pertenecen a una partición así como analiza sus resultados experimentales. Como salida del sub-proceso se generan y empaquetan una o varias piezas de conocimiento.

En la figura 7.2 se muestran los sub-procesos *generador de particiones* y *comparador de replications* que forman parte de este procedimiento, así como los recursos afectados por estos sub-procesos. En esta figura se emplean paquetes UML estereotipados como líneas de ensamble. Una línea de ensamble representa un recurso que es afectado por un

¹Al contar sólo con un experimento no se tiene la certeza si el efecto observado es real o artificial, en esta situación se requiere de una replicación para verificar los resultados. Si la replicación no corrobora los resultados del experimento base se crea una confusión al identificar cuál de los dos resultados si los del experimento o los de la replicación corresponden a una realidad, una tercera replicación ayuda a aclarar esta confusión.

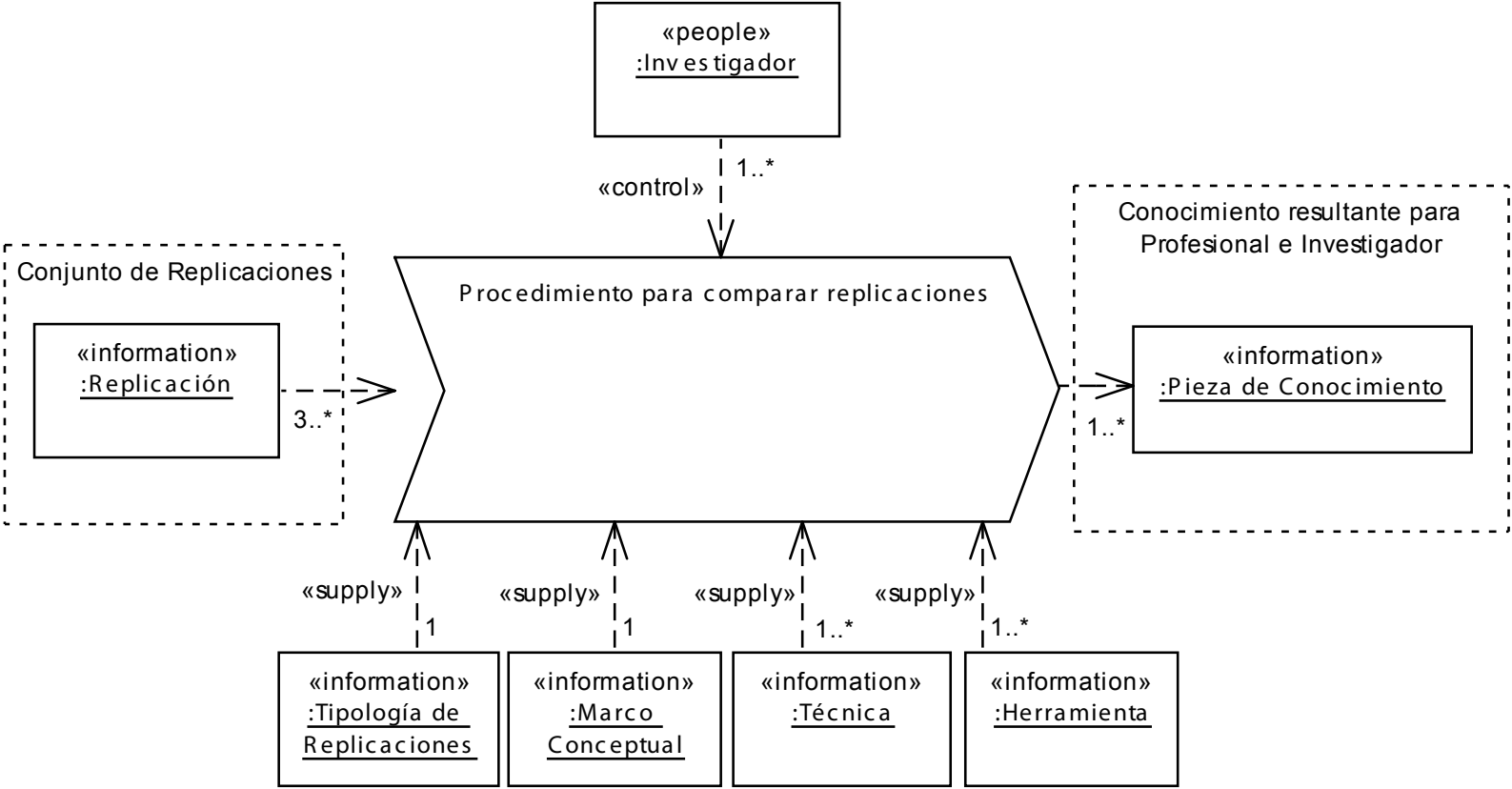


Fig. 7.1: Diagrama general del procedimiento para comparar replicaciones

proceso determinado.

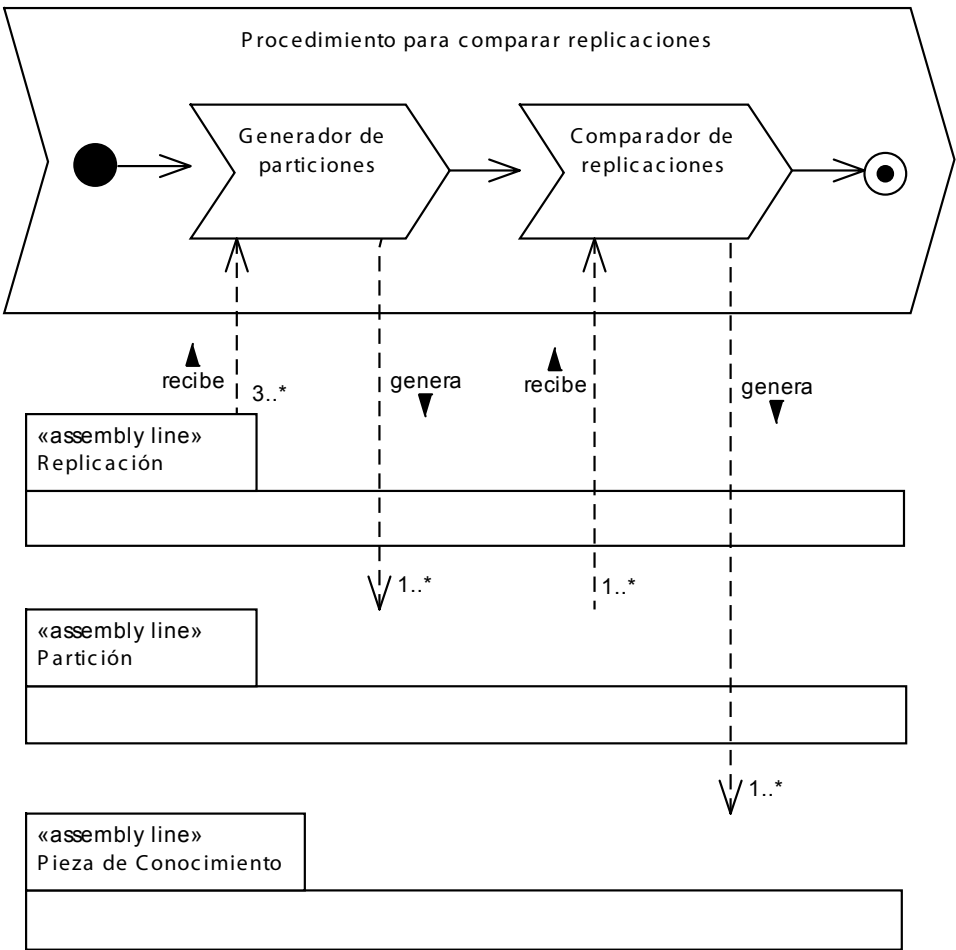


Fig. 7.2: Sub-procesos que forman parte del procedimiento para comparar replicasiones

Una vez presentado de forma general el funcionamiento de este proceso, en los próximos dos capítulos se describe el funcionamiento y las actividades de los sub-procesos *generador de particiones* y *comparador de replicasiones*.

Capítulo 8

Sub-proceso Generador de Particiones

RESUMEN: En este capítulo se presenta el sub-proceso *generador de particiones* que es el primer sub-proceso del procedimiento para comparar replicasiones. Este sub-proceso facilita la generación de particiones de un conjunto de replicasiones, las cuales se usan en el siguiente sub-proceso para llevar a cabo comparaciones entre las replicasiones de una partición.

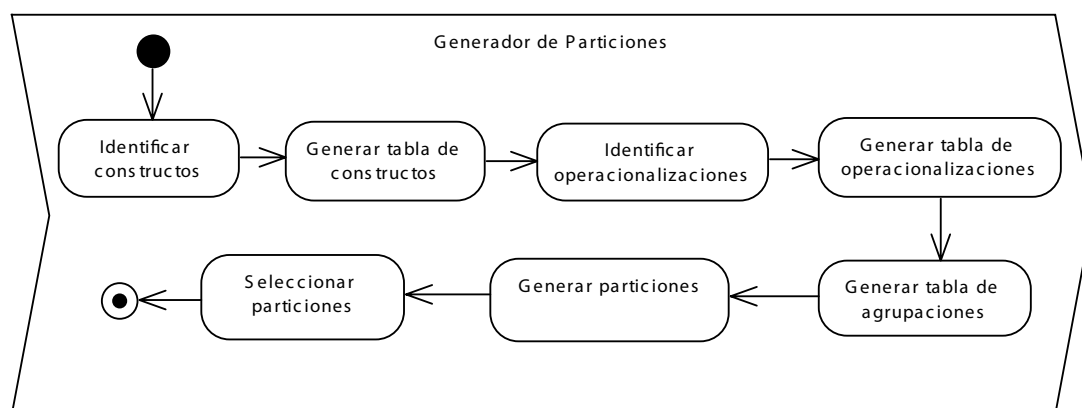
El procedimiento para comparar replicasiones inicia con el sub-proceso *generador de particiones*. Para poder comparar replicasiones y extraer algún conocimiento de ellas, primero es necesario conocer las unidades comparables que tienen en común. El objetivo de este sub-proceso es generar particiones a partir del constructo y operacionalización de efecto así como del par de constructos y operacionalizaciones de causa comunes en las replicasiones. Las particiones resultantes ofrecen al investigador una vista organizada de los constructos y operacionalizaciones estudiados en las replicasiones a comparar.

El sub-proceso *generador de particiones* se compone de las siguientes actividades que se muestran diagramadas en la figura 8.1.

- Identificar constructos
- Generar tabla de constructos
- Identificar operacionalizaciones
- Generar tabla de operacionalizaciones
- Generar tabla de agrupaciones
- Generar particiones
- Seleccionar particiones

8.1 Identificar Constructos

La primera actividad de este sub-proceso consiste en identificar los constructos de causa y efecto de las replicasiones a comparar. Como se describe en el capítulo 6 Los constructos son ideas o conceptos que en un experimento controlado se concretan a través de las operacionalizaciones. Por ejemplo podemos pensar en el funcionamiento de alguna técnica de

Fig. 8.1: Actividades del sub-proceso *generador de particiones*

lectura de código determinada (constructo) y concretarla siguiendo cierto procedimiento para llevarla a acabo en la práctica (operacionalización).

8.2 Generar Tabla de Constructos

Una vez identificados los constructos, la siguiente actividad del sub-proceso consiste en registrarlos en una tabla. Esta tabla se encuentra disponible en el formulario PCR-1 (apéndice D.1). En la tabla 8.1 se muestra un ejemplo de la estructura de esta tabla con información de los constructos identificados en un conjunto de cuatro replicaciones.

Identificador del experimento	Constructos de causa	Constructos de efecto
r1	Técnica de inspección A	Efectividad
	Técnica de inspección B	Eficiencia
	Técnica de inspección C	
r2	Técnica de inspección A	Efectividad
	Técnica de inspección B	
r3	Técnica de inspección A	Efectividad
	Técnica de inspección B	
r4	Técnica de inspección A	Efectividad
	Técnica de inspección B	Eficiencia
	Técnica de inspección C	

Tabla 8.1: Ejemplo de tabla con registro de constructos

8.3 Identificar Operacionalizaciones

La siguiente actividad consiste en identificar las operacionalizaciones de los constructos. Aunque las replicaciones a comparar usen los mismos constructos, éstos pueden ser operacionalizados de distintas maneras. Por ejemplo, la facilidad de uso de un sistema software puede medirse a través de un tipo de medida que indique el tiempo que tarda un usuario en completar una tarea o también puede medirse por el número de errores que comete el usuario al completar una tarea. En este caso el constructo *facilidad de uso* se concreta con dos operacionalizaciones distintas. Un constructo también puede concretarse

con operacionalizaciones similares, i.e. se producen variaciones entre las operacionalizaciones. Por ejemplo, una técnica de inspección de código determinada puede llevarse a cabo siguiendo el procedimiento descrito por cierta fuente, o puede variarse en distintos grados el procedimiento para aplicar la técnica. En un conjunto de replications es posible encontrar tanto operacionalizaciones distintas de un mismo constructo como variaciones en las operacionalizaciones de los constructos.

8.4 Generar Tabla de Operacionalizaciones

Como siguiente actividad, las operacionalizaciones identificadas se registran en una tabla. Esta tabla se encuentra disponible en el formulario PCR-2 (apéndice D.2). En la tabla 8.2 se muestra un ejemplo de esta tabla con la información de las operacionalizaciones identificadas en cuatro replications.

En este ejemplo las replications usan un tipo de medida para medir cada constructo de efecto. La efectividad se mide con la medida A mientras que la eficiencia se mide con la medida B. Como en la replicación r_4 se ha variado la aplicación del procedimiento de las técnicas A y B los tratamientos se identifican con versiones distintas a las replications r_1 , r_2 y r_3 .

Constructos de Causa				Cons. de efecto	
Id. Exp.	Téc. A	Téc. B	Téc. C	Efectividad	Eficiencia
r_1	Téc. Inspección A v1	Téc. Inspección B v1	Téc. Inspección C v1	Medida A	Medida B
r_2	Téc. Inspección A v1	Téc. Inspección B v1		Medida A	
r_3	Téc. Inspección A v1	Téc. Inspección B v1		Medida A	
r_4	Téc. Inspección A v2	Téc. Inspección B v2	Téc. Inspección C v1	Medida A	Medida B

Tabla 8.2: Ejemplo de tabla con registros de operacionalizaciones

8.5 Generar Tabla de Agrupaciones

Una vez que se tiene completa la tabla de operacionalizaciones, la siguiente actividad consiste en organizar las replications en pares de operacionalizaciones y constructos de causa así como por las operacionalizaciones y constructos de efecto. La organización se realiza de acuerdo a la siguiente jerarquía: primero se organizan los constructos de efecto seguido de las operacionalizaciones de efecto, posteriormente se organizan los pares de constructos de causa seguido de las operacionalizaciones de causa. Los pares resultantes se colocan en el formulario PCR-3 (apéndice D.3). En la tabla 8.3 se muestra un ejemplo tras organizar los constructos y operacionalizaciones de cuatro replications.

8.6 Generar Particiones

Una vez que se tiene la tabla de agrupaciones la siguiente actividad de este sub-proceso consiste en generar las particiones. Con la información de la tabla de agrupaciones se generan las particiones organizando las replications de acuerdo a las unidades comparables que tienen en común. En la tabla 8.4 se muestra un ejemplo con las particiones identificadas en cuatro replications.

Las particiones generadas pueden representarse de manera gráfica para facilitar la visualización de las distintas partes comunes que son comparables en un conjunto de

Id. Exp.	Cons. Efecto	Oper. Efecto	Par constructos de Causa	Par de operacionalizaciones de Causa
r1	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A v1 vs. T. Insp. B v1
r1	Efectividad	Medida A	Téc. A vs. Téc. C	T. Insp. A v1 vs. T. Insp. C v1
r1	Efectividad	Medida A	Téc. B vs. Téc. C	T. Insp. B v1 vs. T. Insp. C v1
r1	Eficiencia	Medida B	Téc. A vs. Téc. B	T. Insp. A v1 vs. T. Insp. B v1
r1	Eficiencia	Medida B	Téc. A vs. Téc. C	T. Insp. A v1 vs. T. Insp. C v1
r1	Eficiencia	Medida B	Téc. B vs. Téc. C	T. Insp. B v1 vs. T. Insp. C v1
r2	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A v1 vs. T. Insp. B v1
r3	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A v1 vs. T. Insp. B v1
r4	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A v2 vs. T. Insp. B v2
r4	Efectividad	Medida A	Téc. A vs. Téc. C	T. Insp. A v2 vs. T. Insp. C v1
r4	Efectividad	Medida A	Téc. B vs. Téc. C	T. Insp. B v2 vs. T. Insp. C v1
r4	Eficiencia	Medida B	Téc. A vs. Téc. B	T. Insp. A v2 vs. T. Insp. B v2
r4	Eficiencia	Medida B	Téc. A vs. Téc. C	T. Insp. A v2 vs. T. Insp. C v1
r4	Eficiencia	Medida B	Téc. B vs. Téc. C	T. Insp. B v2 vs. T. Insp. C v1

Tabla 8.3: Ejemplo de pares de constructos y operacionalizaciones identificados

replicaciones. En la figura 8.2 se muestran las particiones de la tabla 8.4 representadas de manera gráfica. Esta representación se ha hecho empleando mapas estructurados. En el apéndice B se describe el funcionamiento de esta técnica para usarla como parte de esta actividad.

En esta figura, las replicaciones $r1$, $r2$ y $r3$ pertenecen a una partición cuyo par de operacionalizaciones de causa son técnica de inspección A versión 1 vs. técnica de inspección B versión 1. Otro tipo de partición que se puede generar se presenta cuando en lugar de agrupar la replicaciones por pares de operacionalizaciones de causa éstas se agrupan por el par de constructos de causa. Esta partición se genera para saber (a través del sub-proceso *comparador de replicaciones*) si se presenta una generalización de los pares de operacionalizaciones al observar un efecto similar en las replicaciones que pertenecen a un mismo par de constructos de causa. Este segundo tipo de partición se recomienda realizar una vez que se ha ejecutado el sub-proceso *comparador de replicaciones* con el primer tipo de partición i.e. aquel que contiene las comparaciones a nivel de operacionalizaciones de causa y efecto. Dependiendo de los resultados que se obtengan al ejecutar el *comparador de replicaciones* con este primer tipo de particiones, se determina si es viable realizar la generación de particiones a nivel de constructos con el fin de generalizar los resultados experimentales a nivel de operacionalizaciones.

Part.	Cons. Efecto	Oper. Efecto	Par Cons. Causa	Par Oper. Causa	Id. Exp.	Exps. por Partición
01	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A v1 vs. T. Insp. B v1	r1, r2, r3	3
02	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A v2 vs. T. Insp. B v2	r4	1
03	Efectividad	Medida A	Téc. A vs. Téc. C	T. Insp. A v1 vs. T. Insp. C v1	r1	1
04	Efectividad	Medida A	Téc. A vs. Téc. C	T. Insp. A v2 vs. T. Insp. C v1	r4	1
05	Efectividad	Medida A	Téc. B vs. Téc. C	T. Insp. B v1 vs. T. Insp. C v1	r1	1
06	Efectividad	Medida A	Téc. B vs. Téc. C	T. Insp. B v2 vs. T. Insp. C v1	r4	1
07	Eficiencia	Medida B	Téc. A vs. Téc. B	T. Insp. A v1 vs. T. Insp. B v1	r1	1
08	Eficiencia	Medida B	Téc. A vs. Téc. B	T. Insp. A v2 vs. T. Insp. B v2	r4	1
09	Eficiencia	Medida B	Téc. A vs. Téc. C	T. Insp. A v1 vs. T. Insp. C v1	r1	1
10	Eficiencia	Medida B	Téc. A vs. Téc. C	T. Insp. A v2 vs. T. Insp. C v1	r4	1
11	Eficiencia	Medida B	Téc. B vs. Téc. C	T. Insp. B v1 vs. T. Insp. C v1	r1	1
12	Eficiencia	Medida B	Téc. B vs. Téc. C	T. Insp. B v2 vs. T. Insp. C v1	r4	1

Tabla 8.4: Ejemplo de particiones resultantes en cuatro replicaciones

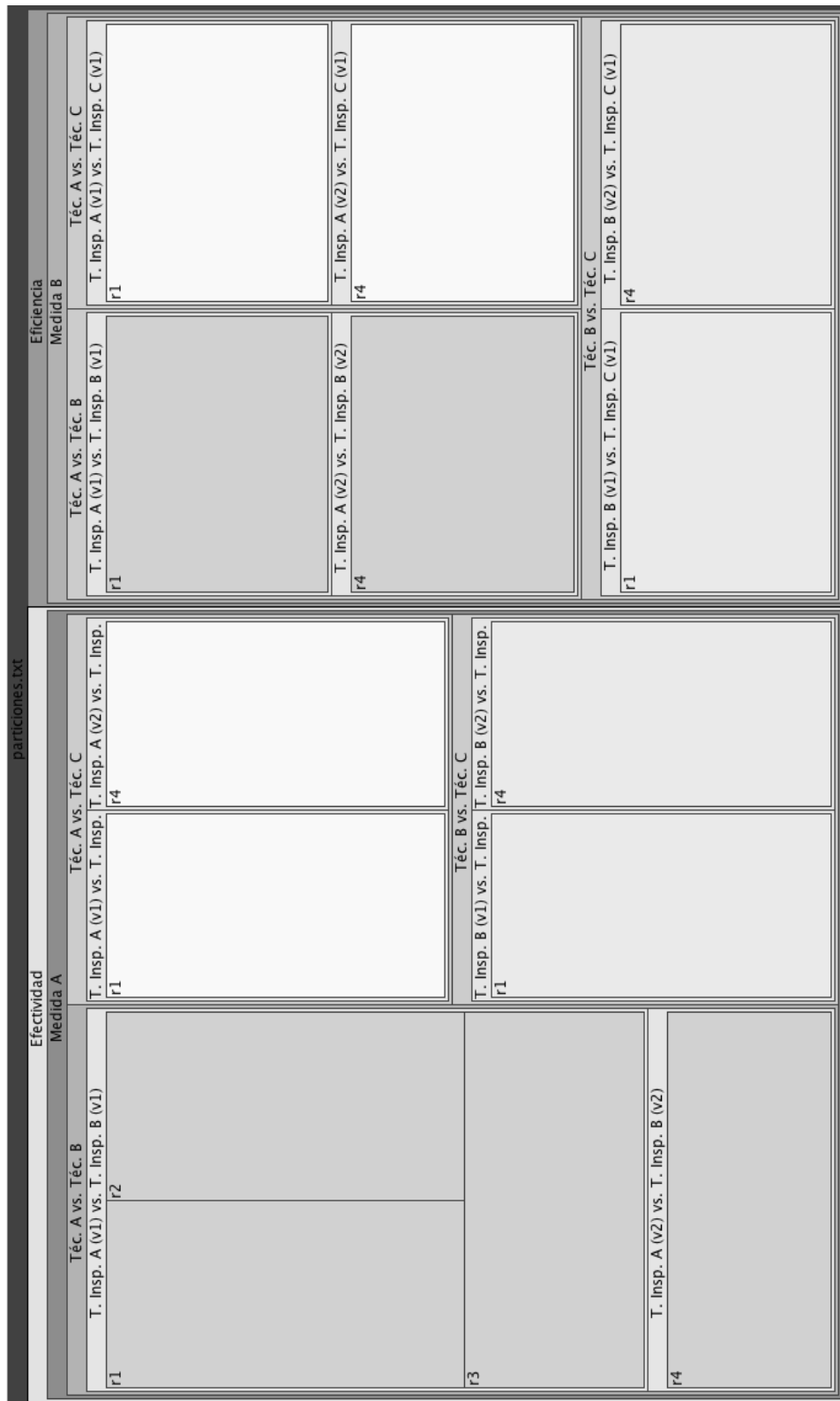


Fig. 8.2: Particiones representadas como un mapa estructurado

8.7 Seleccionar Particiones

En esta última actividad del sub-proceso se seleccionan las particiones a utilizar en el siguiente sub-proceso (*comparador de replicaciones*). En esta actividad se seleccionan sólo las particiones en las que se pueden realizar las comparaciones. Para realizar las comparaciones se recomienda que cada partición contenga al menos tres replicaciones. Esto con el fin de 1) reducir la confusión de posibles resultados experimentales contradictorios y 2) generar conocimiento con cierto nivel de confiabilidad. Por ejemplo, de las particiones identificadas en la tabla 8.4 de la actividad anterior, la partición 01 puede usarse para realizar comparaciones, ya que esta partición agrupa tres replicaciones. En la tabla 8.5 se muestra esta partición.

Part.	Cons. Efecto	Oper. Efecto	Par Cons. Causa	Par Oper. Causa	Id. Exp.	Exps. por Partición
1	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A v1 vs. T. Insp. B v1	r1, r2, r3	3

Tabla 8.5: Partición valida para usar en el *comparador de replicaciones*

Una vez descritas las actividades que forman parte del sub-proceso *generador de particiones*, en el siguiente capítulo se describe el funcionamiento del sub-proceso *comparador de replicaciones*.

Capítulo 9

Sub-proceso Comparador de Replicaciones

RESUMEN: En este capítulo se describe el segundo sub-proceso del procedimiento para comparar replications. El *comparador de replications* permite al investigador comparar replications en una partición respecto a sus situaciones experimentales y a sus resultados experimentales. Estas comparaciones permiten conocer el grado de verificación logrado en las replications de una partición así como empaquetar el conocimiento obtenido.

Una vez que se tiene alguna partición, este segundo sub-proceso realiza comparaciones entre las replications de una partición para obtener algún conocimiento. El objetivo de este sub-proceso consiste en generar conocimiento a partir de la comparación de cambios realizados en las replications de una partición respecto a los elementos de la situación experimental así como de la comparación de replications a nivel de resultados experimentales.

El primer tipo de comparación se realiza para conocer los cambios que se han realizado en las replications de una partición, este tipo de comparación facilita al investigador conocer los diferentes elementos de la situación experimental que han variado en las replications de una partición. La identificación de cambios facilita conocer la similitud que hay entre las replications de una partición. El segundo tipo de comparación se realiza para conocer qué tan semejantes son los resultados de las replications en una partición. Con los resultados de estas dos comparaciones es posible conocer el tipo de verificación que se logra cuando los resultados en una partición son consistentes, identificar variables moderadoras cuando los resultados no son consistentes, así como empaquetar el conocimiento obtenido en una o más piezas.

9.1 Estructura del Comparador de Replicaciones

El *comparador de replications* se compone de una actividad y tres sub-procesos que son:

- Generar tabla maestra
- Sub-proceso Identificación de Similitudes
- Sub-proceso Análisis de Resultados Experimentales
- Sub-proceso Empaquetado de Piezas de Conocimiento

En la figura 9.1 se muestra la actividad y los sub-procesos que forman parte del *comparador de replicaciones* así como los recursos con los que interactúa.

Como entrada se recibe una o varias particiones previamente seleccionadas, a partir de estas particiones se genera la tabla maestra que contiene descripciones de los elementos de la situación experimental de las replicaciones que conforman las particiones seleccionadas. En el sub-proceso *identificación de similitudes* se recibe una o varias particiones para examinar los diferentes cambios que hay en las replicaciones de cada partición, como salida se obtiene una tabla con los cambios identificados en las particiones así como dos tipos de visualizaciones que representan la similitud entre las replicaciones respecto a los cambios en los elementos de la situación experimental. El sub-proceso *análisis de resultados experimentales* recibe como entrada una o varias particiones para examinar los resultados de las replicaciones que pertenecen a una partición. Como resultado se obtiene una tabla con los resultados de las comparaciones así como visualizaciones que muestran las comparaciones de los resultados de las replicaciones de cada partición. El último sub-proceso del *empaquetado de piezas de conocimiento* se encarga de empaquetar el conocimiento obtenido en piezas que puedan ser útiles tanto para el profesional como para el investigador. Este sub-proceso recibe como entrada una o varias tablas de cambios generadas en el sub-proceso *identificación de similitudes* así como una o varias tablas con los resultados de las comparaciones del sub-proceso *análisis de resultados experimentales*. A Partir de estas tablas, en este último sub-proceso se empaqueta el conocimiento obtenido de las distintas comparaciones. A continuación se describe la actividad *generar tabla maestra* de este primer sub-proceso.

9.2 Generar Tabla Maestra

Una vez elegidas las particiones para realizar comparaciones y obtener conocimiento de ellas, en esta actividad se genera la tabla maestra que contiene los elementos de la situación experimental de las replicaciones pertenecientes a las particiones seleccionadas. El objetivo de la tabla maestra es registrar los distintos elementos de la situación experimental presentes en las replicaciones de las particiones seleccionadas.

Como primer paso se identifican las replicaciones que hay en las particiones seleccionadas. Por cada una de las replicaciones se identifican y extraen los elementos de la situación experimental. Los elementos identificados en cada replicación pueden registrarse en el formulario PCR-4 (apéndice D.4). Esta tarea conlleva cierto esfuerzo ya que puede dificultarse la identificación de los elementos de la situación experimental en los artículos, documentos o fuentes donde se describa la replicación.

Si bien esta tarea puede realizarse por un investigador, es posible involucrar a otros investigadores. Esta tarea puede llevarse a cabo en una o varias sesiones, dependiendo del número de particiones y el número de replicaciones que haya en cada partición. Cuando participan varios investigadores en esta tarea, cada investigador puede trabajar en la identificación de elementos de una replicación y la sesiones de trabajo pueden aprovecharse para resolver dudas y consolidar una visión común de los términos usados por los distintos autores de las replicaciones a comparar. Al existir una retroalimentación en cada sesión, se produce un refinamiento y homogenización en la información extraída de las replicaciones que conforman las particiones seleccionadas.

Una vez registrados los elementos de la situación experimental de cada replicación, éstos se colocan en una tabla maestra. Esta tabla se encuentra disponible en el formulario PCR-5 (apéndice D.5). En la tabla 9.1 se muestra un ejemplo con un fragmento de la tabla

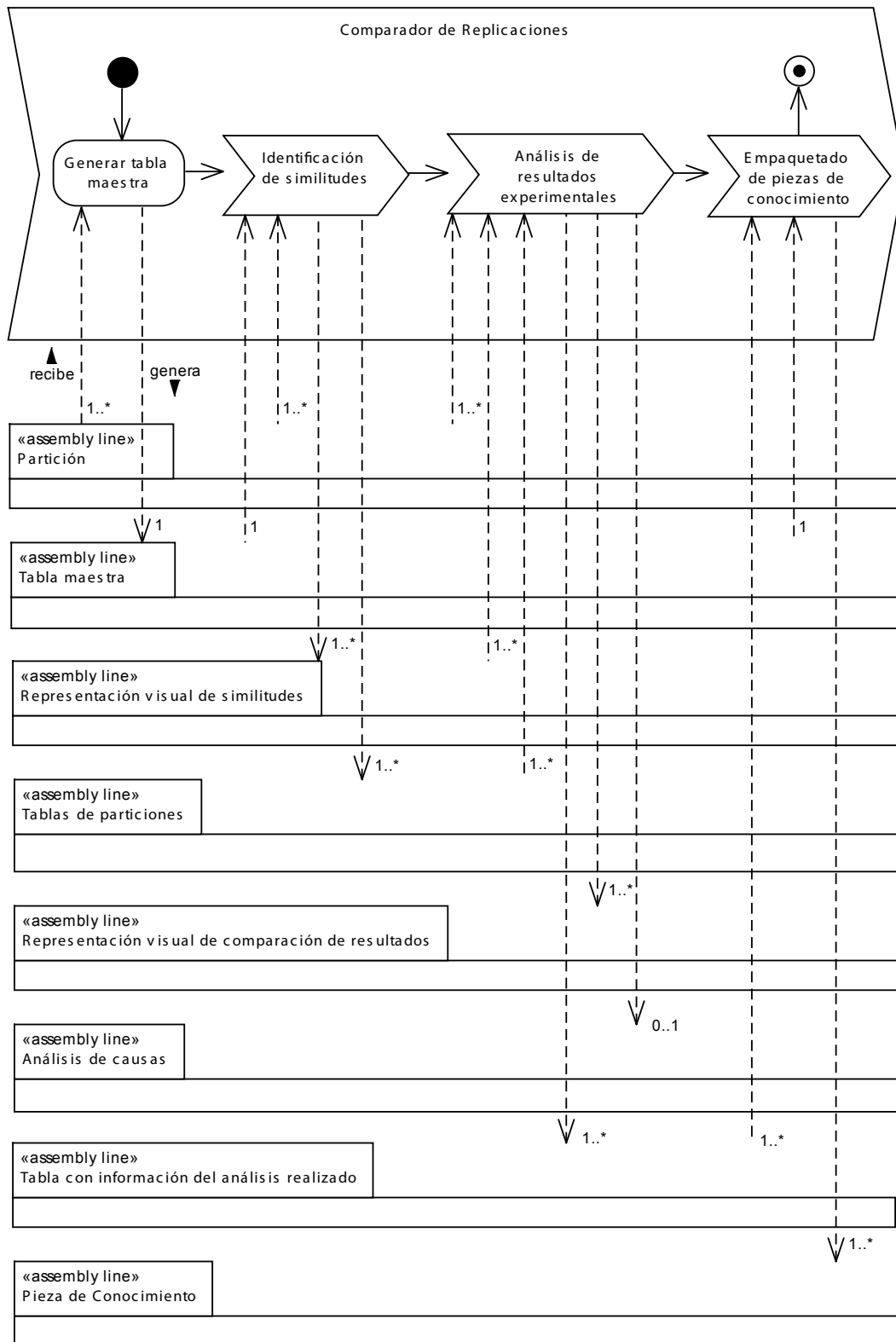


Fig. 9.1: Sub-procesos e interacción del *comparador de replicaciones*

maestra referente a la sección del protocolo experimental. En este ejemplo se seleccionaron dos particiones que en conjunto suman siete replicaciones.

Esta actividad del sub-proceso se realiza sólo una vez, ya que esta tabla contiene la información del conjunto de particiones a utilizar en el resto de sub-procesos del *comparador de replicaciones*.

9.3 Sub-proceso Identificación de Similitudes

Este es el primer sub-proceso que forma parte del *comparador de replicaciones*. El objetivo de este sub-proceso es examinar los cambios y la similitud que hay entre las replicaciones de una partición. La similitud de las replicaciones se examina de acuerdo a los elementos de la situación experimental.

Como entrada, este sub-proceso recibe replicaciones pertenecientes a una partición. A partir de la partición suministrada, en este sub-proceso se extrae la información de los distintos elementos de la situación experimental de cada replicación. Con la información compilada en la tabla maestra y con el uso de técnicas estadísticas multivariantes se generan representaciones visuales que indican la similitud entre las replicaciones de la partición. Como salida se obtienen varias tablas con información de los diferentes elementos de la situación experimental que se han utilizado en las replicaciones así como representaciones visuales que indican la similitud de las replicaciones de una partición. Los resultados de este sub-proceso permiten conocer los cambios y la similitud que hay en las replicaciones de cada partición. Este sub-proceso se compone de las siguientes actividades que junto al sub-proceso se muestran en la figura 9.2.

- Generar tablas de particiones
- Identificar similitudes

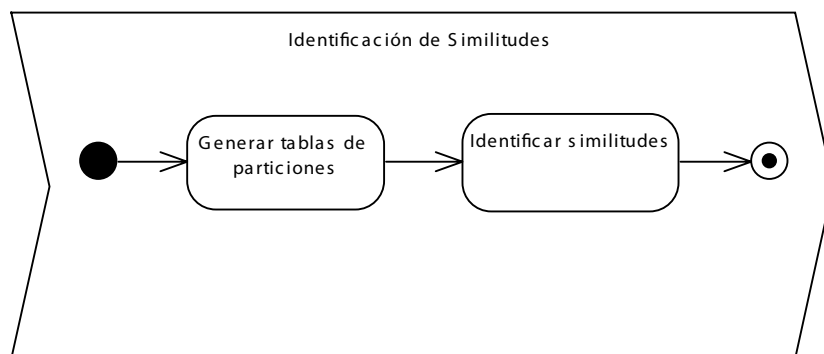


Fig. 9.2: Actividades del sub-proceso *identificación de similitudes*

9.3.1 Generar Tablas de Particiones

La tabla maestra generada en la actividad anterior de este sub-proceso sirve de referente para crear las tablas de particiones. Estas tablas contienen los elementos de la situación experimental de las replicaciones que conforman cada partición. Con la información de cada tabla es posible conocer el nivel de similitud que hay en las replicaciones de una partición ya que es posible identificar los cambios realizados en las replicaciones.

		Protocolo experimental						
		Partición 1				Partición 2		
Elemento principal	Sub-elemento / Replicación	r1	r2	r3	r4	r5	r6	r7
Análisis	Técnica de análisis de datos	ANOVA mixto	ANOVA mixto	ANOVA mixto	ANOVA de una vía	ANOVA de una vía	ANOVA mixto	ANOVA factorial
Diseño	Diseño experimental	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto
	Orden de aplicación de los tratamientos	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden
	Combinación del Entrenamiento con las sesiones	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial
	Tratamientos por sesión	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa
	Sesiones requeridas por sujeto para aplicar tratamiento	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión
	Tareas por sujeto por sesión	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa
Instrumentos	Instrumentos o equipo de medición	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)
Objetos experimentales	Programas	P1, P2, P3	P1, P2, P4	P1, P3, P4	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee
	Faltas (Instancias)	Instancia A	Instancia B	Instancia C	Instancia D	Instancia E	Instancia F	Instancia G y H
Materiales	Formularios del experimento	Formularios v1	Formularios v1	Formularios v1	Formularios v2	Formularios v2	Formularios v2	Formularios v2
	Instrucciones del experimento	Instrucciones v1	Instrucciones v1	Instrucciones v1	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2
	Guías suplementarias	Guías v1	Guías v1	Guías v1	Guías v2	Guías v2	Guías v2	Guías v2

Tabla 9.1: Fragmento de tabla maestra con replicaciones pertenecientes a dos particiones

La primera actividad de este sub-proceso consiste en generar a partir de la información de la tabla maestra las tablas de particiones. Las tablas resultantes se usan en la siguiente actividad de este sub-proceso para calcular los niveles de similitud que hay en las replicas de cada partición. Ya que la siguiente actividad tiene cierto grado de automatización se requiere que en esta actividad las particiones se especifiquen de acuerdo a un formato específico. En el apéndice D.6 se encuentra el formulario PCR-6 que contiene la estructura para registrar la información de las tablas de particiones. En la tabla 9.2 se muestra un ejemplo con un fragmento de la tabla de partición referente al protocolo experimental. En este ejemplo la tabla de partición contiene tres replicas pertenecientes a una partición.

Estructura	Elemento Principal	Sub Elemento	ID	Categorías	r1	r2	r3
Protocolo experimental	Análisis	Técnica de análisis de datos	3.1.1	ANOVA 1 vía	X		X
			3.1.1	ANOVA mixto		X	
	Diseño	Diseño Experimental	3.2.1	Diseño Aleatorio	X		X
			3.2.1	Diseño Cruzado		X	
		Orden de aplicación de los tratamientos	3.2.2	Aleatorio	X		X
			3.2.2	2 tipos de orden		X	
		Combinación del entrenamiento con las sesiones	3.2.3	Secuencial	X	X	X
		Tratamientos por sesión	3.2.4	3 técnicas, 1 programa	X		
			3.2.4	2 técnicas, 1 programa		X	X
		Sesiones requeridas por sujeto para aplicar tratamiento	3.2.5	1 sesión	X	X	X
		Tareas por sujeto por sesión	3.2.6	1 técnica, 1 programa	X		X
			3.2.6	2 técnicas, 1 programa		X	
	Instrumentos	Instrumentos o equipos de medición	3.3.1	Ordenador, folios, bolígrafo	X	X	X
	Objetos experimentales	Programas (Códigos fuente)	3.4.1	Programa A	X	X	X
			3.4.2	Instancia de Faltas A	X	X	X
	Materiales	Formularios del experimento	3.5.1	Formularios v1	X	X	X
		Instrucciones del experimento	3.5.2	Instrucciones v1	X	X	X
		Guías suplementarias	3.5.3	Guías v1	X	X	X

Tabla 9.2: Sección de la tabla de partición referente al protocolo experimental

A diferencia de la estructura de la tabla maestra, en las tablas de particiones las categorías que describen el contenido de los (sub)elementos ¹ de la situación experimental se transforman en valores dicotómicos, de tal manera que las categorías de cada (sub)elemento se marcan con el símbolo “X” para indicar su presencia en las replicas de la partición. En estas tablas también se añade una columna que contiene un código numérico que identifica de forma jerárquica los (sub)elementos de la situación experimental.

9.3.2 Identificar Similitudes

Una vez que se ha rellenado la tabla de particiones con los distintos elementos de la situación experimental, la siguiente actividad consiste en determinar y representar el nivel de

¹Son sub-elementos cuando se refiere a las secciones de protocolo, operacionalizaciones de constructos y poblaciones, y elementos cuando se refiere a las secciones de experimentadores y sitio.

similitud que hay entre las replications que componen una partición. Para determinar la similitud de situaciones experimentales entre las replications se calcula una matriz numérica cuyos valores indican el nivel de similitud que hay entre las replications de una partición. A partir de esta matriz se usan técnicas estadísticas multivariantes como son el análisis de conglomerados jerárquico AHC (*Agglomerative Hierarchical Clustering*) y el análisis de escalado multidimensional MDS (*Multidimensional Scaling*, Escalado Multidimensional) para examinar de manera visual las similitudes que hay en las replications de una partición. En el apéndice B se describe el funcionamiento de estas técnicas.

La similitud entre replications se determina a partir de la información registrada en la tabla de particiones. Para determinar los niveles de similitud es necesario establecer algún tipo de medida de proximidad que cuantifique el nivel de similitud que hay entre las replications. Las medidas de proximidad se usan para calcular distancias o (dis)similitudes que hay entre varios objetos, en este caso, calcular similitudes entre replications. En el apéndice B se describen varias medidas de proximidad así como la medida de proximidad usada para la calcular la similitud entre las replications de una partición.

Las similitudes se calculan a partir de los diferentes pares de replications que hay en una partición. Las similitudes resultantes de cada par de replications se organizan en una matriz simétrica de $N \times N$ que se le conoce como matriz de similitudes. En el apéndice B se describe en detalle como calcular esta matriz.

En la tabla 9.3 se muestra un ejemplo de una matriz con las similitudes de tres replications pertenecientes a una partición. Las similitudes están representadas como distancias donde valores cercanos a 0 indican mayor similitud mientras que valores cercanos a 1 indican mayor dis-similitud. En este ejemplo las replications $r2$ y $r3$ se asemejan más entre sí.

	r1	r2	r3
r1	0		
r2	0.6831	0	
r3	0.6633	0.3872	0

Tabla 9.3: Matriz de similitud de una partición con tres replications

Una vez que se dispone la matriz de similitudes de una partición, es posible representar las similitudes de manera gráfica. A través de técnicas estadísticas multivariantes como son el análisis de conglomerados jerárquico AHC o el análisis de escalado multidimensional MDS que usan una matriz de similitudes, es posible examinar visualmente las similitudes que hay en las replications de una partición respecto a los cambios identificados en los elementos de las situaciones experimentales. En la figura 9.3 se muestran dos tipos de diagramas que representan las similitudes en una partición compuesta de tres replications. En 9.3a se muestra un dendrograma como resultado de aplicar AHC y en 9.3b se muestra un plano de coordenadas tras aplicar MDS.

Para facilitar la elaboración de cálculos que implica esta actividad, calcular las distintas matrices de similitud así como aplicar las técnicas AHC y MDS, se ha realizado un procedimiento de automatización para identificar las similitudes en las replications de una partición. En el apéndice B se detalla esta automatización.

Al realizar las diferentes actividades que componen este sub-proceso, el investigador adquiere el conocimiento respecto a los cambios realizados y el nivel de similitud que hay entre las replications que conforman una partición. Los distintos cambios identificados en las replications de las particiones se usan en el siguiente sub-proceso.

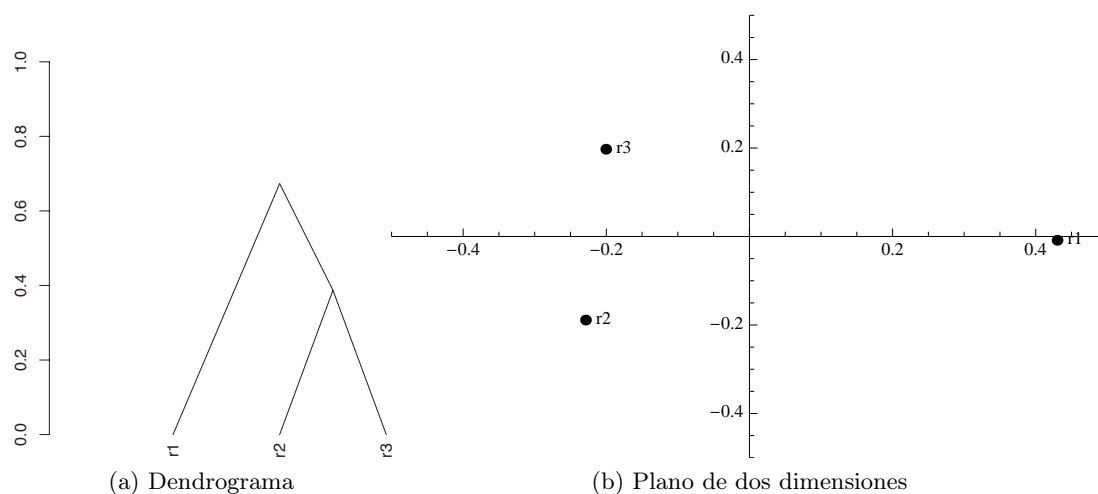


Fig. 9.3: Dos tipos de gráfico que representan las similitudes entre tres replicasiones

9.4 Sub-proceso Análisis de Resultados Experimentales

Una vez ejecutado el sub-proceso *identificación de similitudes*, el siguiente sub-proceso del *comparador de replicasiones* consiste en realizar el análisis de resultados experimentales. El objetivo de este sub-proceso es comparar los resultados de las replicasiones de una partición para conocer el grado de verificación obtenido en caso de observar resultados consistentes o conocer posibles elementos de la situación experimental que se cree influyen en los resultados.

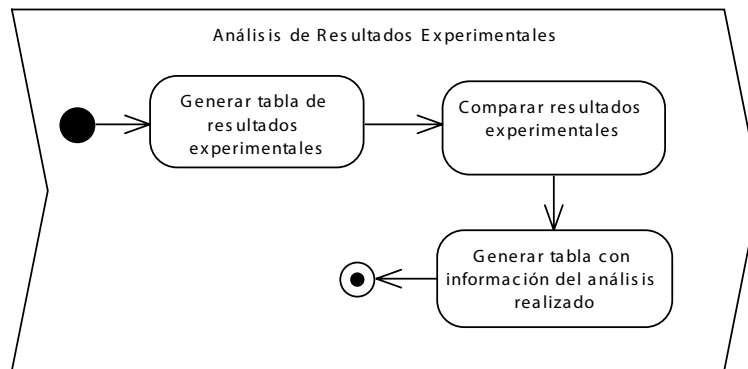
Como entrada de este sub-proceso se recibe una partición así como las tablas y los gráficos generados en el sub-proceso anterior. A partir de los recursos suministrados, en este sub-proceso se comparan los resultados de las replicasiones de una partición empleando técnicas estadísticas de control de procesos SPC (*Statistical Process Control*) así como de agregación de resultados². Como salida se obtienen varias representaciones visuales de las comparaciones resultantes así como una tabla con información de los resultados de las comparaciones. Los resultados de este sub-proceso permiten al investigador conocer el grado de verificación logrado cuando se obtienen resultados consistentes en la partición o en caso contrario, conocer posibles elementos de la situación experimental que se cree influyen en los resultados. En la figura 9.4 se muestra el sub-proceso *análisis de resultados experimentales* que se compone de las siguientes actividades:

- Generar tabla de resultados experimentales
- Comparar resultados experimentales
- Generar tabla con información del análisis realizado

9.4.1 Generar Tabla de Resultados Experimentales

La primera actividad de este sub-proceso consiste en extraer la información relacionada a los resultados experimentales de cada una de las replicasiones de la partición en la que se está trabajando. De acuerdo a la estructura de una partición, ésta se compone por el constructo y operacionalización de efecto de un par de constructos y operacionalizaciones

²En el apéndice B se describe en detalle el funcionamiento de estas dos técnicas.

Fig. 9.4: Actividades del sub-proceso *análisis de resultados experimentales*

de causa. Por ejemplo, la métrica o variable de respuesta de un par de tratamientos. En este sentido la información a extraer de cada replicación consta de los siguientes elementos:

- El tamaño de la muestra utilizada n en cada tratamiento
- El valor promedio observado de cada tratamiento \bar{x}
- La desviación estándar s del par de tratamientos

La información extraída de las replicaciones de la partición se coloca en la tabla del formulario PCR-7 descrito en el apéndice D.7. En la tabla 9.4 se muestra un ejemplo con la estructura principal del formulario PCR-7. En este ejemplo se tienen los resultados experimentales de tres replicaciones que evalúan la efectividad representada con la métrica A de dos técnicas. Estas replicaciones conforman una partición determinada.

Id exp.	Técnica A			Técnica B		
	Tamaño (n_1)	Métrica A (\bar{x}_1)	Desv. est. (s_1)	Tamaño (n_2)	Métrica A (\bar{x}_2)	Desv. est. (s_2)
r1	45	34.1	12.7	45	28.1	11.5
r2	39	29.7	22.1	39	22.6	25.8
r3	43	32.9	17.5	43	37.5	21.2

Tabla 9.4: Resultados experimentales de tres replicaciones que pertenecen a una partición

El promedio y la desviación estándar que aparecen en la tabla 9.4 están representados en una escala porcentual. Por ejemplo, en la técnica A de la replicación *r1* se observa una efectividad respecto a la métrica A del 34.1% respecto a la técnica B donde se observa una efectividad del 28.1%.

9.4.2 Comparar Resultados Experimentales

Una vez que se disponen los resultados experimentales de la partición seleccionada, la siguiente actividad de este sub-proceso consiste en realizar las comparaciones de estos resultados. La comparación de los resultados se lleva a cabo a través del uso de técnicas estadísticas de control de procesos y de técnicas estadísticas de agregación de resultados i.e. meta-análisis³.

³Ambas técnicas se describen en profundidad en el apéndice B.

Como técnica estadística de control de procesos se emplean diagramas de gráficos de control. Este tipo de gráfico se usa con el fin de examinar la variabilidad en los tratamientos de las replicaciones que componen una partición. Por otra parte se utiliza meta-análisis para combinar resultados. A través del meta-análisis es posible conocer la distribución de los tamaños de efecto observados en varias replicaciones, así como estimar un tamaño de efecto global a partir de los tamaños de efecto observados en un conjunto de replicaciones. En este caso, el meta-análisis se usa para examinar la distribución de los tamaños de efecto observados en las replicaciones de una partición, así como para estimar un tamaño de efecto global.

Los resultados obtenidos en los gráficos de control y el meta-análisis junto con los gráficos generados en el sub-proceso *identificación de similitudes* permiten al investigador conocer el tipo de verificación logrado en caso de obtener resultados reproducibles o en caso contrario, conocer las causas que se cree afectan los resultados de una partición. En esta actividad se realizan las siguientes tareas:

- Generación de gráficos de control
- Generación de meta-análisis
- Examinar resultados experimentales y similitudes
- Análisis de causas de resultados inconsistentes

9.4.2.1 Generación de Gráficos de Control

La primera tarea de esta actividad consiste en representar la variabilidad de cada tratamiento aplicado en las replicaciones de la partición. Los gráficos de control facilitan examinar la distribución de un conjunto de observaciones en un periodo de tiempo determinado. En el caso de este sub-proceso, cada observación representa el valor promedio o la magnitud de un tratamiento aplicado en una replicación. En el apéndice B se describe el funcionamiento de los gráficos de control.

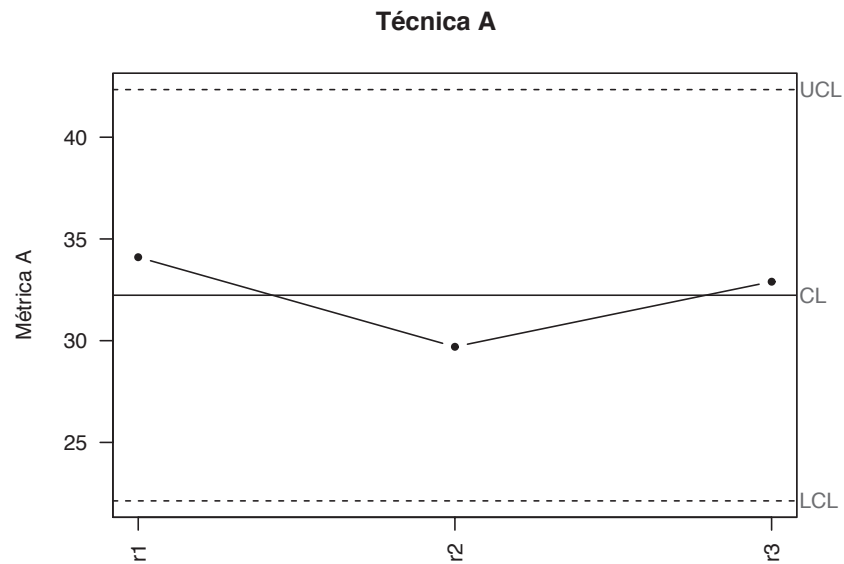
Usando como referencia la información de los dos tratamientos que aparece en la tabla de resultados experimentales (tabla 9.4) descrita en la actividad anterior, se generan los gráficos de control por cada tratamiento. En la figura 9.5 se muestra un ejemplo de dos gráficos de control de dos tratamientos en este caso técnica A y B aplicadas en las replicaciones de una partición.

Como se observa en esta figura, las observaciones de los tratamientos se encuentran dentro de los límites de control, por lo que las variaciones en las magnitudes de ambos tratamientos pueden atribuirse a una variabilidad natural.

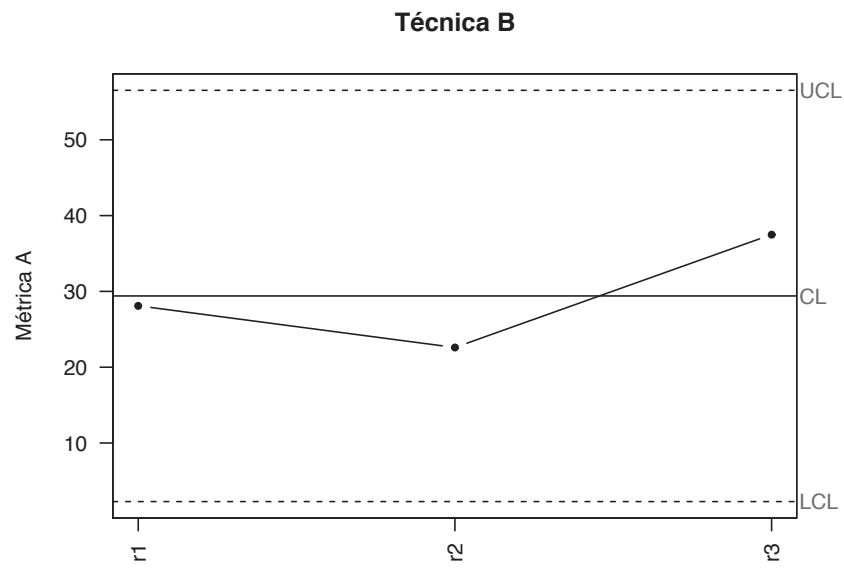
9.4.2.2 Generación de Meta-análisis

Una vez examinada la variabilidad individual de los tratamientos, la siguiente tarea de esta actividad consiste en obtener los tamaños de efecto de las replicaciones de la partición seleccionada. Para esto es necesario calcular un tamaño de efecto estimado a partir de la información de los pares de tratamientos de la partición. Este paso se realiza usando técnicas de agregación de resultados i.e. meta-análisis. En el apéndice B se describe detalladamente el funcionamiento de esta técnica.

En la figura 9.6 se muestra un ejemplo del meta-análisis generado en una partición. En este ejemplo se usa la tabla de resultados experimentales descrita en la tabla 9.4 de la actividad anterior.



(a) Ejemplo de tratamiento representado por la técnica A



(b) Ejemplo de tratamiento representado por la técnica B

Fig. 9.5: Gráficos de control de dos tratamientos

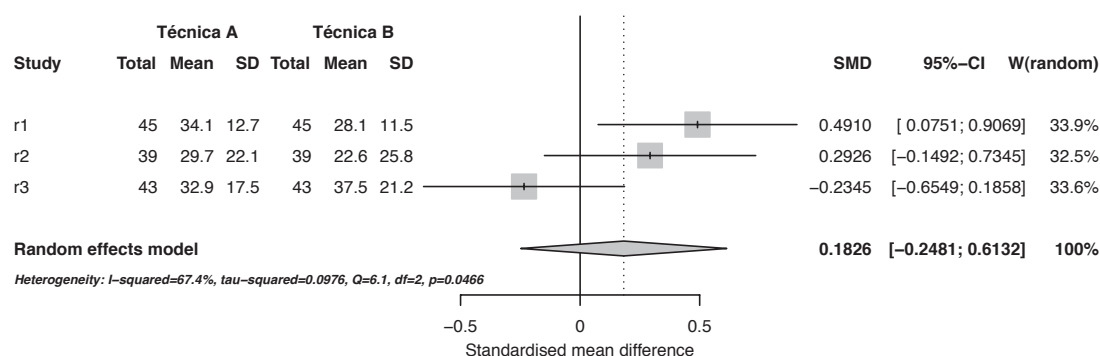


Fig. 9.6: Ejemplo de meta-análisis

De acuerdo a los resultados del meta-análisis, se observa que los tamaños de efecto estimados en las replicaciones de la partición son heterogéneos⁴, se observa que la replicación $r3$ contiene un efecto negativo que a diferencia del resto de replicaciones indica que la técnica B es ligeramente más efectiva⁵ que la técnica A.

Para facilitar la elaboración de cálculos que implica esta actividad, calcular los promedios, límites de control de cada tratamiento, generar los gráficos de control así como realizar los distintos cálculos para estimar los tamaños de efecto individual, global, intervalos de confianza y calcular los niveles de heterogeneidad, se ha realizado un procedimiento de automatización de cálculos para el análisis de resultados experimentales de una partición. En el apéndice B se detalla esta automatización.

9.4.2.3 Examinar Resultados Experimentales y Similitudes

Una vez que se tienen los resultados de los gráficos de control y meta-análisis, la siguiente tarea de esta actividad consiste en examinar estos resultados junto con los resultados del sub-proceso *identificación de similitudes* para determinar el tipo de verificación que se logra en las replicaciones de una partición.

En los gráficos de control y el meta-análisis se examina la consistencia de los resultados experimentales i.e. la variabilidad en los gráficos de control así como la heterogeneidad en el meta-análisis, mientras que en la identificación de similitudes se examinan los cambios y similitudes en las replicaciones de una partición respecto a los elementos de la situación experimental. De acuerdo al tipo de resultados experimentales obtenidos y a los cambios identificados en las situaciones experimentales de las replicaciones de una partición, pueden presentarse cuatro posibles combinaciones que en la tabla 9.5 se describen.

Tipo de resultados experimentales	Tipo de cambio en situaciones experimentales	Confianza de resultados	Verificación de resultados
Resultados consistentes	Sin Cambios	Confiable	Se confirma resultado
Resultados consistentes	Con Cambios	Confiable	Se confirma resultado y hay ganancia de conocimiento
Resultados inconsistentes	Sin Cambios	No Confiable	-
Resultados inconsistentes	Con Cambios	No Confiable	-

Tabla 9.5: Combinaciones de resultados que pueden obtenerse

⁴El resultado de la prueba de heterogeneidad es significativo, de acuerdo al valor establecido $\alpha = 0.05$, $p = 0.0466$; por otra parte los valores de T^2 e I^2 son elevados.

⁵Aunque la diferencia no es significativa ya que el tamaño del efecto real puede ser 0.

La primera combinación se presenta cuando los resultados tanto de los gráficos de control como del meta-análisis son consistentes i.e. las observaciones en los gráficos de control se encuentran dentro de los límites de control y los resultados del meta-análisis indican nula o poca heterogeneidad, además en los resultados de la identificación de similitudes no se observan cambios en los elementos de la situación experimental.

En esta combinación el conocimiento que se obtiene es confiable ya que los resultados son consistentes entre las replications de la partición, sin embargo al mantener sin cambios los elementos de la situación experimental, el efecto global observado podría tener menor poder de confirmación de resultados en caso de tener mayor conocimiento sobre las condiciones que influyen en los resultados experimentales⁶.

La segunda combinación se presenta cuando en las replications de la partición los resultados experimentales tanto de los gráficos de control como del meta-análisis son consistentes y los resultados de la identificación de similitudes se observan cambios en los elementos de la situación experimental. En esta situación el conocimiento que se obtiene es confiable y el efecto global observado tiene mayor poder de confirmación de resultados. Al variar las replications y confirmar un mismo resultado, es mucho más probable que el efecto global observado se deba a una realidad y no a la configuración del protocolo empleado en las replications.

La tercera combinación se presenta cuando se observan inconsistencias en los resultados experimentales ya sea por una variabilidad fuera de límites en uno o ambos tratamientos y/o una heterogeneidad en la distribución de los tamaños de efecto estimados, además, no se observan cambios en los elementos de la situación experimental de las replications de una partición. En esta combinación, se observa que alguna condición influye en los resultados de los gráficos de control y/o meta-análisis, sin embargo al no observar cambios en la situación experimental es más difícil encontrar la causa de los resultados inconsistentes. Una interpretación de este resultado puede ser que en el sub-proceso *identificación de similitudes* no se identificaron apropiadamente los cambios en los elementos de la situación experimental. Otra interpretación puede ser que los resultados son así de variables. En este tipo de resultado, el conocimiento que se obtiene no es confiable debido a la inconsistencia de resultados.

La cuarta combinación se presenta cuando se observan inconsistencias en los resultados experimentales y también se observan cambios en los elementos de la situación experimental. En esta combinación lo más probable es que alguno de los cambios identificados en la situación experimental influya en los resultados experimentales. En este caso el conocimiento que se obtiene no es confiable debido a la contradicción en la confirmación de los resultados.

Cuando se presentan resultados inconsistentes en alguna de las particiones, el investigador decide intentar obtener algún conocimiento a partir de este tipo de resultados o detenerse en este punto del *comparador de replications* y seleccionar otra partición. No obstante, se recomienda continuar e intentar obtener conocimiento ante la presencia de resultados inconsistentes con el fin de identificar condiciones experimentales que se cree influyen en los resultados y estudiarlas posteriormente con mayor profundidad. A continuación se describen las situaciones donde es posible obtener conocimiento en caso de obtener resultados inconsistentes en una partición.

Cuando se observan resultados inconsistentes en una partición se recomienda examinar la distribución de las magnitudes de los tratamientos en los gráficos de control así como la

⁶De acuerdo a Collins (Collins, 1985) cuando en una disciplina se tiene conocimiento maduro, el poder de confirmación de experimentos similares (replications) es menor.

distribución de los tamaños de efecto en el meta-análisis. Tras examinar las distribuciones se puede presentar alguna de las siguientes situaciones:

1. En las distribuciones se observan grupos de resultados.
2. En las distribuciones se observa ausencia de grupos de resultados (no se observa algún patrón aparente).

La primera situación se presenta cuando se observan grupos de resultados ya sea en los gráficos de control y/o en el diagrama de bosque (generado tras aplicar meta-análisis). En la figura 9.7 se muestra un ejemplo de varios diagramas de bosque donde se tienen diferentes agrupaciones de resultados⁷. En las figuras 9.7a y 9.7b se muestran dos agrupaciones de resultados mientras que en la figura 9.7c tres agrupaciones.

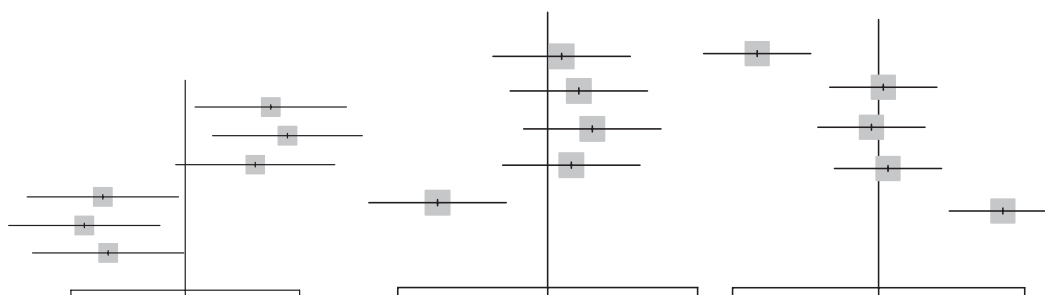


Fig. 9.7: Ejemplos de agrupaciones de resultados en un diagrama de bosque con resultados heterogéneos

En la segunda situación es difícil observar agrupaciones de resultados. En esta situación los resultados se distribuyen sin algún patrón aparente. En la figura 9.8 se muestra un diagrama de bosque bajo esta situación.

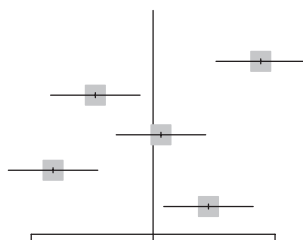


Fig. 9.8: Tamaños de efecto distribuidos a lo largo del diagrama de bosque

Cuando se presenta la primera situación es posible obtener algún conocimiento. En esta situación se identifican grupos formados por al menos dos resultados⁸ que parezcan ser consistentes entre sí, el resto de resultados se apartan para su posterior análisis. Por ejemplo, asumiendo que las inconsistencias se deben a una heterogeneidad en el meta-análisis (las observaciones en los gráficos de control se encuentran dentro de sus límites de control). En la figura 9.9 se muestran varios diagramas de bosque donde se identifican algunos grupos de resultados consistentes.

⁷Las agrupaciones de resultados serían similares para los gráficos de control.

⁸Esto con el fin de tener al menos la verificación de un resultado.

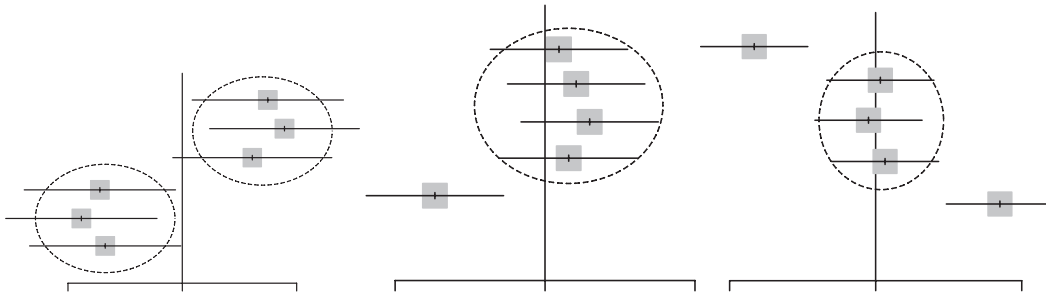


Fig. 9.9: Ejemplos de agrupaciones de resultados consistentes

Con los grupos de resultados consistentes identificados es posible obtener algún conocimiento⁹, sin embargo, al observar previamente inconsistencias en los resultados de la partición se recomienda llevar a cabo un análisis de causas.

El análisis de causas ayuda al investigador a rastrear e identificar posibles elementos de la situación experimental que se cree influyeron en los resultados. Los resultados de este análisis sirven también para definir hipótesis que pueden ser testadas en replications futuras i.e. descubrir nuevas condiciones o factores explicativos que clarifiquen las diferencias en los resultados experimentales.

Por último, cuando se presenta la segunda situación es difícil obtener conocimiento de la partición debido a la dispersión de los resultados. Sin embargo en esta situación también se puede llevar a cabo un análisis de causas.

Un ejemplo de esta tarea se presenta a continuación. Al examinar los resultados de los gráficos de control (figura 9.10), meta-análisis (figura 9.11) e identificación de similitudes (figura 9.12), se observa que los resultados en los gráficos de control se encuentran dentro de los límites de control, sin embargo los resultados en el meta-análisis indican heterogeneidad. Por otra parte, la identificación de similitudes muestra cambios en algunos elementos de la situación experimental.

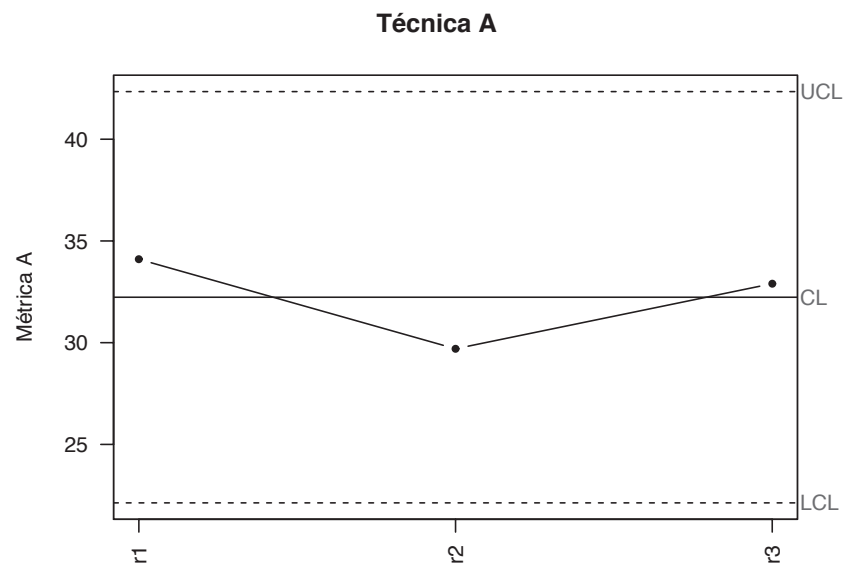
De acuerdo a los cuatro combinaciones de resultados que se pueden obtener, en este ejemplo se presenta la cuarta combinación. Se observan resultados inconsistentes debido a la heterogeneidad identificada y se tienen cambios en los elementos de la situación experimental.

En este punto, se decide intentar obtener algún conocimiento de esta partición o descartarla. Para seguir con este ejemplo, se decide continuar e intentar obtener conocimiento a partir de los resultados inconsistentes de la partición.

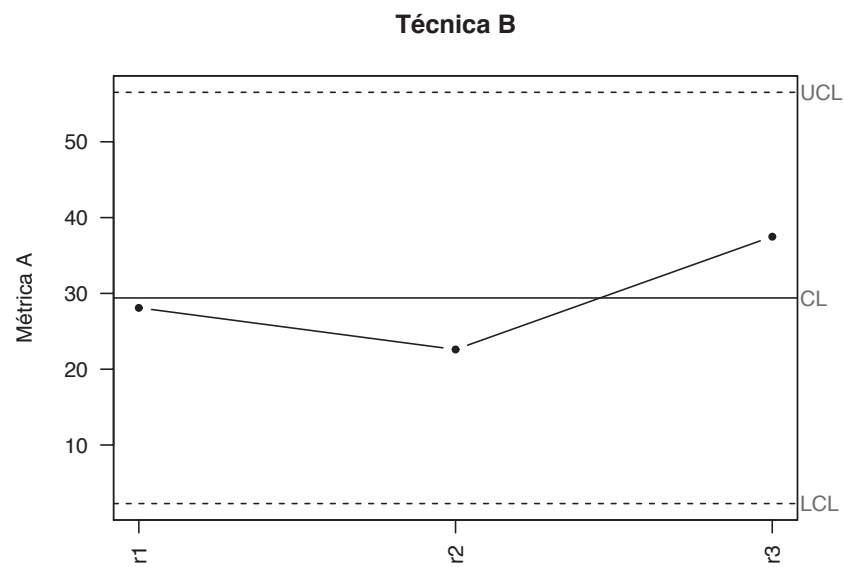
Tras reconocer las inconsistencias en los resultados, a continuación se examina las distribuciones de observaciones en los gráficos de control y las distribuciones de tamaños de efecto en el meta-análisis. Las observaciones en los gráficos de control se encuentran dentro de los límites de control, sin embargo en el meta-análisis se observa una heterogeneidad en los resultados. En el diagrama de bosque se observan dos grupos de resultados. Las replications $r1$ y $r2$ forman el primer grupo mientras que $r3$ el segundo. En este caso se seleccionan las replications $r1$ y $r2$ cuyos resultados se asemejan más. Estas dos replications se eligen para obtener conocimiento de ellas.

Una vez seleccionadas las replications del grupo de resultados, se realizan de nuevo los gráficos de control y meta-análisis para verificar que los resultados sean consistentes

⁹En el caso del meta-análisis cuando se apartan estudios para obtener resultados consistentes se le denomina análisis de sensibilidad (Borenstein et al., 2009).



(a) Ejemplo de tratamiento representado por la técnica A



(b) Ejemplo de tratamiento representado por la técnica B

Fig. 9.10: Gráficos de control de dos tratamientos

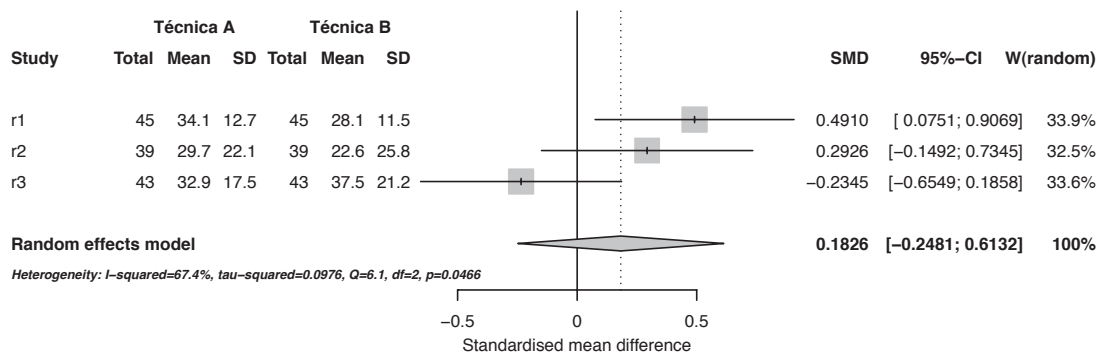


Fig. 9.11: Meta-análisis de la partición seleccionada

entre sí. En las figuras 9.13 y 9.14 se muestran los resultados de estos nuevos cálculos.

como se observa en los resultados de los gráficos de control y meta-análisis, los resultados son suficientemente consistentes para poder obtener conocimiento confiable a partir de las replicaciones $r1$ y $r2$. Sin embargo, al apartar replicaciones de la partición para obtener resultados consistentes se recomienda llevar a cabo un análisis de causas con el fin de identificar posibles causas de inconsistencias en los resultados de la partición.

9.4.2.4 Análisis de Causas de Resultados Inconsistentes

El análisis de causas puede llevarse a cabo por uno o varios investigadores. El análisis consiste en identificar en las tablas de partición (Formulario PCR-6 D.6) los cambios realizados en las replicaciones de la partición y marcarlos con algún símbolo en los diagramas de causa y efecto. Una vez que se tienen marcados los cambios en los diagramas, se examina cada cambio y se discute si alguno de los cambios es responsable de influir en los resultados experimentales. Los cambios identificados pueden colocarse en la tabla del formulario PCR-8 que se encuentra disponible en el apéndice D.8. Los posibles cambios que se creen causan la inconsistencia de resultados se usan para generar nuevas hipótesis y testarlas en replicaciones futuras.

Puede presentarse el caso que en los resultados de la identificación de similitudes no se observen cambios en los elementos de la situación experimental y por el contrario se observen inconsistencias en los resultados i.e. la tercera combinación de resultados antes descrita. En esta situación es más difícil identificar las causas de inconsistencia en los resultados ya que no se identificaron cambios en la situación experimental que pudieron influir en los resultados. En esta situación se puede realizar el análisis de causas con el fin de tratar de identificar el o los elementos de la situación experimental que pudieron influir en los resultados.

Como se mencionó anteriormente, al encontrar resultados inconsistentes el investigador decide si continúa o no con las actividades de este sub-proceso. Si se decide continuar se recomienda llevar a cabo el análisis de causas.

En las figuras 9.15 y 9.16 se presentan ejemplos de diagramas de causa y efecto con cambios identificados en elementos de la situación experimental que se cree influye en los resultados. Los cambios se representan por un número que se encuentra debajo del (sub)elemento que ha variado. Este número indica las veces que se ha variado el (sub)elemento.

Los cambios identificados en los (sub)elementos de la situación experimental se mues-

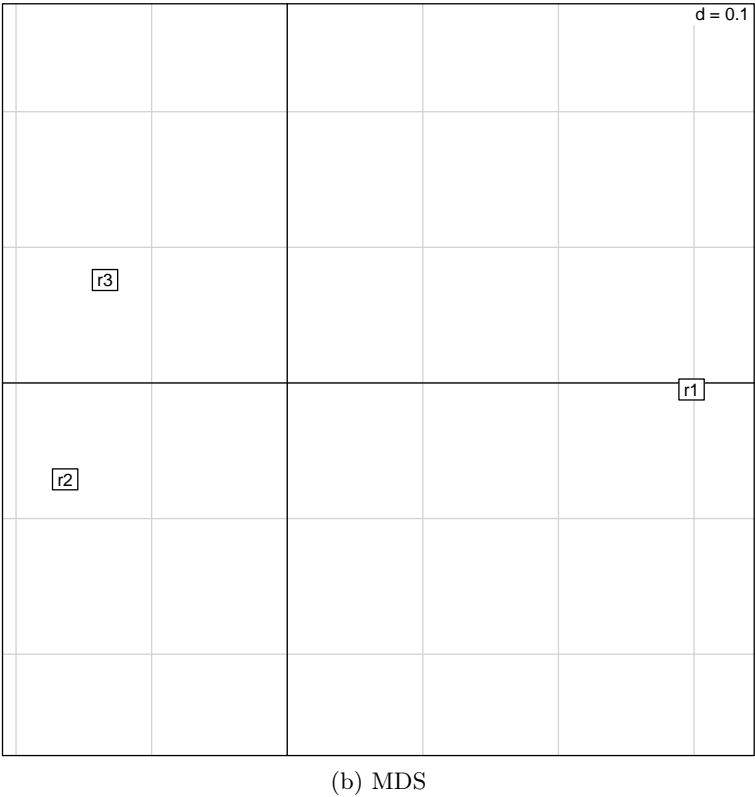
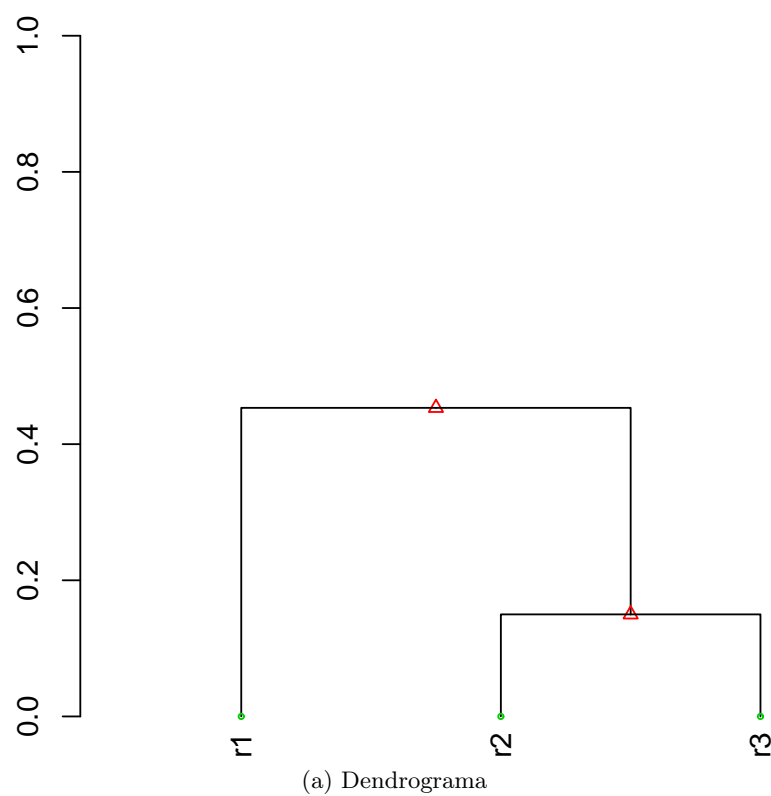
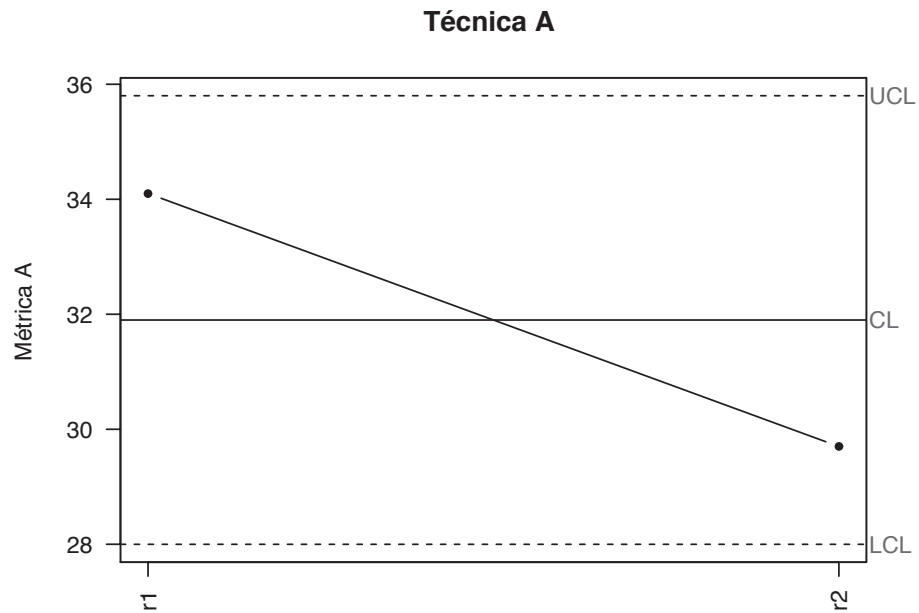
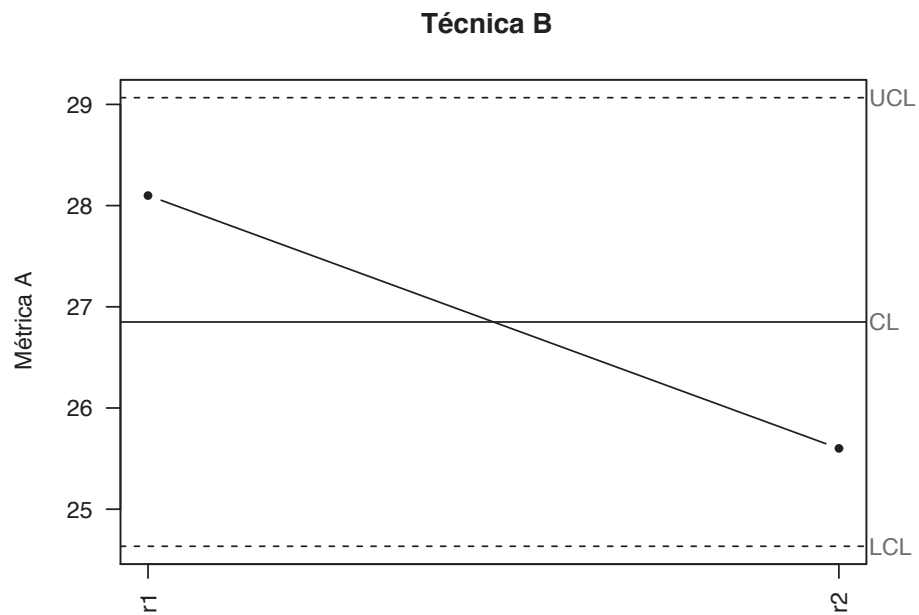


Fig. 9.12: Representación de similitudes resultante



(a) Ejemplo de tratamiento representado por la técnica A



(b) Ejemplo de tratamiento representado por la técnica B

Fig. 9.13: Gráficos de control tras apartar la replicación r_3

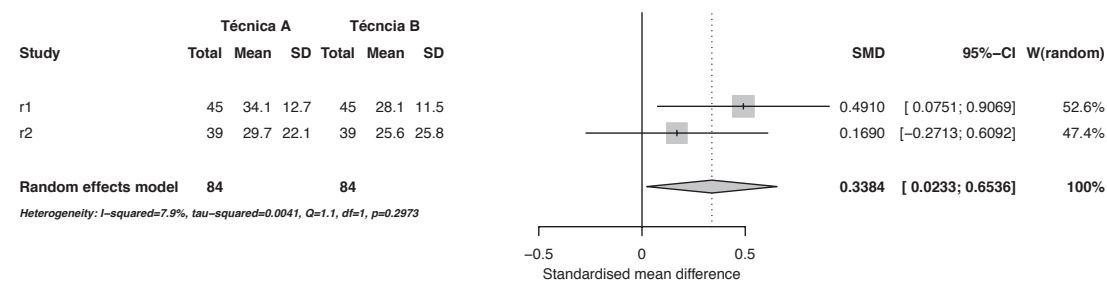


Fig. 9.14: Meta-análisis de la partición seleccionada sin replicación $r3$

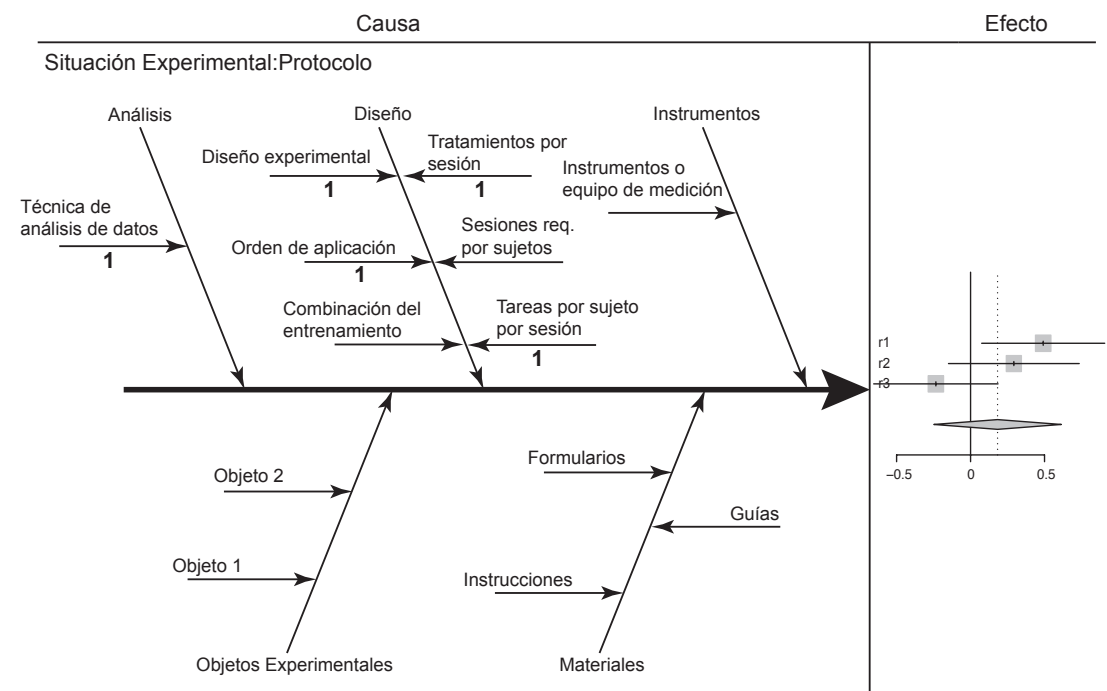


Fig. 9.15: Diagrama de causa y efecto con número de cambios identificados en protocolo experimental

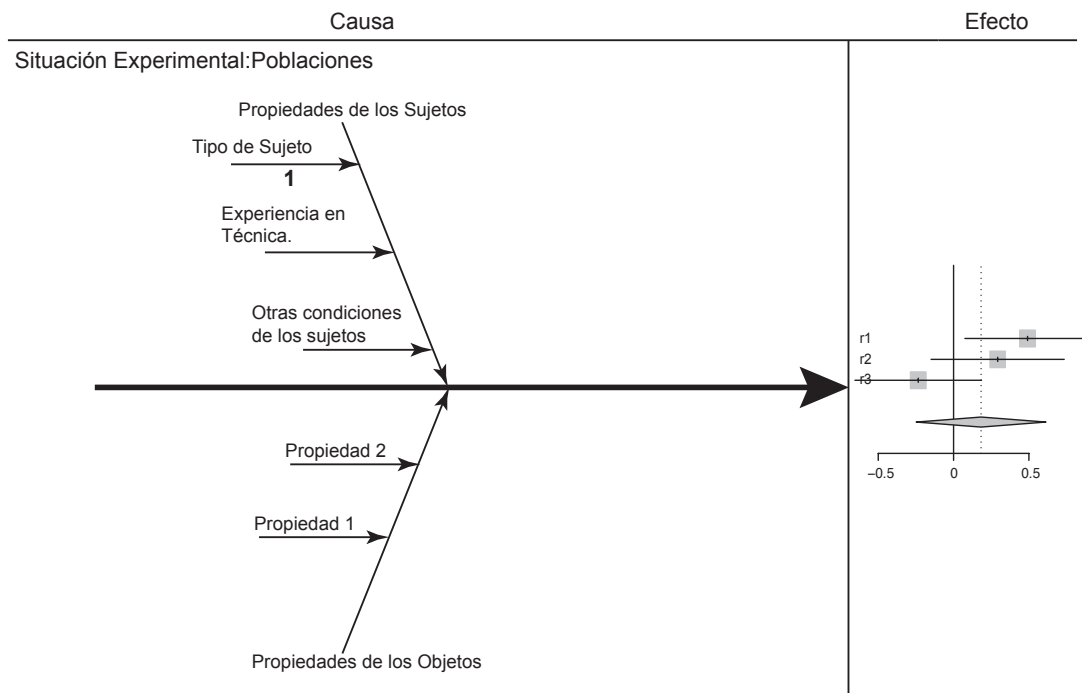


Fig. 9.16: Diagrama de causa y efecto con el número de cambios identificados en las poblaciones

tran en la tabla 9.6. A continuación se examina cada uno de los cambios y se discute si alguno de los cambios es probablemente responsable de la variación en los resultados experimentales.

Un ejemplo de discusión de este análisis se describe a continuación. Comparando los cambios y los resultados entre las replicas $r1$ y $r2$, se observa que los cambios en $r2$ no influyeron en sus resultados y éstos se asemejan a los resultados de $r1$. Respecto a la replicación $r3$, los cambios relacionados con el entrenador, monitor y tipo de sujeto contienen un tipo de cambio que no es común en las replicas $r1$ o $r2$. De estos tres tipos de variación existe mayor probabilidad de que el tipo de sujeto haya influido en los resultados de la replicación $r3$. A diferencia de las replicas $r1$ y $r2$, en $r3$ se emplearon sujetos con un nivel de experiencia mayor. Tal vez el contar con una mayor experiencia los sujetos detectan más faltas que los sujetos de pregrado empleados en las replicas $r1$ y $r2$. Bajo este supuesto, se elige al tipo de sujeto como posible causa de variación en la replicación $r3$.

9.4.3 Generar Tabla con Información del Análisis Realizado

Una vez que se ha realizado la comparación de los resultados experimentales, la última actividad de este sub-proceso consiste en generar una tabla con la información de las comparaciones de resultados realizadas en la actividad anterior, esta información se usa para obtener conocimiento y empaquetarlo en una o más piezas. La estructura de esta tabla se encuentra disponible en el formulario PCR-9 descrito en el apéndice D.9. Esta tabla se divide en cuatro secciones que describen la información de la partición, información del tipo de resultados obtenidos en las comparaciones, información de resultados consistentes

Sección	Elemento	Subelemento	Categorías	r1	r2	r3
Experimentadores	Diseñador		Experimentador A	X		
			Experimentador C		X	X
	Entrenador		Experimentador A	X		
			Experimentador C		X	X
			Experimentador D			X
	Monitor		Experimentador A	X		
			Experimentador B	X		
			Experimentador C		X	X
			Experimentador D			X
	Medición		Experimentador B	X		
			Experimentador C		X	X
	Analista		Experimentador B	X		
			Experimentador C		X	X
Sitio	Sitio		Sitio A	X		
			Sitio B		X	X
Protocolo experimental	Análisis	Técnica de análisis de datos	ANOVA 1 vía	X		X
			ANOVA mixto		X	
	Diseño	Diseño Experimental	Diseño Aleatorio	X		X
			Diseño Cruzado		X	
		Orden de aplicación de los tratamientos	Aleatorio	X		X
			2 tipos de orden		X	
		Tratamientos por sesión	3 técnicas, 1 programa	X		
			2 técnicas, 1 programa		X	X
Poblaciones	Propiedades de los sujetos	Tipo de sujeto	1 técnica, 1 programa	X		X
			2 técnicas, 1 programa		X	
			Pregrado	X	X	
			Grado (Máster)			X

Tabla 9.6: Cambios identificados en elementos de la situación experimental

e información del análisis de causas.

La primera sección identificada como *información de la partición* se compone de los siguientes elementos:

- Id. Partición. Que representa el identificador de la partición.
- Partición. Que es una descripción de la partición utilizada en este sub-proceso.
- Replicaciones de la partición. Representa el número de replicaciones de la partición así como sus identificadores.

La segunda sección de la tabla *Información del tipo de resultados obtenidos* se compone de los elementos:

- Resultados del análisis de similitudes. En este elemento se describe si en el sub-proceso *identificación de similitudes* se observaron cambios o no, en los elementos de la situación experimental.
- Tipo de Resultados encontrados. Indica si se observaron o no inconsistencias en los resultados de la partición seleccionada.
- Información de los resultados. En este elemento se describen los resultados tanto de los gráficos de control como del meta-análisis. Respecto a los gráficos de control se especifican los límites de control superior e inferior, el promedio y las observaciones. En el meta-análisis se especifican los resultados de la prueba de heterogeneidad (estadístico Q y su valor p) así como los estadísticos I^2 y T^2 .
- Tipo de inconsistencia observada. En este elemento se describen las características de la inconsistencia, por ejemplo, presencia de heterogeneidad debido a los tamaños de efecto observados en una de las replicaciones de la partición.

La tercera sección de la tabla contiene información de los resultados consistentes, los elementos que conforman esta sección son:

- Replicaciones consideradas para la pieza de conocimiento. En este elemento se indican las replicaciones de la partición que se usarán para generar la pieza de conocimiento. Se especifica el número de replicaciones seleccionadas y sus identificadores.
- Comparación del par de tratamientos. En este elemento se indica si el efecto del tratamiento A es mayor, menor o igual que el tamaño del efecto del tratamiento B.
- Tipo de efecto global. Este elemento indica si el tamaño de efecto global estimado es o no significativo.
- Tamaño del efecto global. Indica el tamaño del efecto global estimado así como los intervalos de confianza estimados para el tamaño del efecto global verdadero.

Por último, la sección de análisis de causas contiene el elemento:

- Nuevas hipótesis de trabajo. En caso de realizar el análisis de causas, en este elemento se colocan las condiciones que se cree son la causa de las inconsistencias observadas en los resultados de las replicaciones de la la partición. Con estas condiciones se construyen nuevas hipótesis de trabajo para ser testadas en futuras replicaciones.

En la tabla 9.7 se muestra un ejemplo de la estructura principal del formulario PCR-9 (disponible en apéndice D.9) con la información de los resultados obtenidos en la actividad anterior. En la primera sección de la tabla se muestra la información de la partición usada, en este caso la partición 01. En la segunda sección se muestra el tipo de resultados obtenidos tras realizar las comparaciones, en este ejemplo se presenta heterogeneidad de resultados en el meta-análisis ocasionada en gran medida por la replicación $r3$. Al apartar esta replicación y realizar de nuevo los gráficos de control y el meta-análisis se obtienen resultados consistentes. En la siguiente sección se muestra la información de resultados consistentes, donde la técnica A es significativamente más efectiva que la técnica B. En la última sección de la tabla se muestran las condiciones que se cree influyeron en los resultados de la partición, de acuerdo al análisis de causas, se cree que la experiencia de los sujetos utilizada en la replicación $r3$ influyó en el tamaño del efecto estimado.

9.5 Sub-proceso Empaquetado de Piezas de Conocimiento

El empaquetado de piezas de conocimiento es el último sub-proceso del *comparador de replicaciones*. El objetivo de este sub-proceso es empaquetar el conocimiento obtenido tras comparar las replicaciones de una partición. Este conocimiento es empaquetado en piezas que puedan ser de interés para el profesional o para el investigador.

Como entrada, este sub-proceso recibe la tabla del formulario PCR-9 (D.9) con la información del análisis realizado en el sub-proceso *análisis de resultados experimentales* y las tablas de particiones del formulario PCR-6 (D.6) generadas en el sub-proceso *identificación de similitudes*. A partir de la información de las tablas de particiones, en este sub-proceso se genera una tabla con información de los cambios realizados en la situación experimental de las replicaciones de la partición cuyos resultados son consistentes. Con esta tabla y con la tabla resultante del análisis de resultados experimentales, este sub-proceso genera como salida, piezas de conocimiento tanto para el profesional como para el investigador. El sub-proceso *empaquetado de piezas de conocimiento* se compone de las siguientes actividades que junto al sub-proceso se muestran en la figura 9.17.

Información de la Partición	
Id. Partición:	01
Partición:	Efectividad/ Métrica A/ Téc. A vs Téc. B/ T. Insp. A v1 vs. T. Insp. B v1
Replicaciones de la partición:	3 (r1, r2, r3)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes
Información de los resultados:	Q=6.1, p=0.0466, I ² =67.4%, T ² = 0.0976 Téc. A: 32.23% [22.12; 42.34] obs: 34.1, 29.7, 32.9 Téc. B: 29.4% [2.27; 56.53] obs: 28.1, 22.6, 37.5
Tipo de inconsistencia observada:	Heterogeneidad en meta-análisis debido a experimento r3
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	2 (r1, r2)
Comparación del par de tratamientos:	T. Insp. A es mayor que T. Insp. B
Tipo de efecto global:	Significativo
Tamaño del efecto global:	0.3384 (Pequeño) [0.0233; 0.6536]
Análisis de Causas	
Nuevas hipótesis de trabajo:	La experiencia de los sujetos influye en los resultados de estas dos técnicas

Tabla 9.7: Información de las comparaciones de resultados realizadas en las replicaciones de la partición

- Generar tabla de variaciones en situación experimental
- Generar pieza de conocimiento para profesional
- Generar pieza de conocimiento para investigador

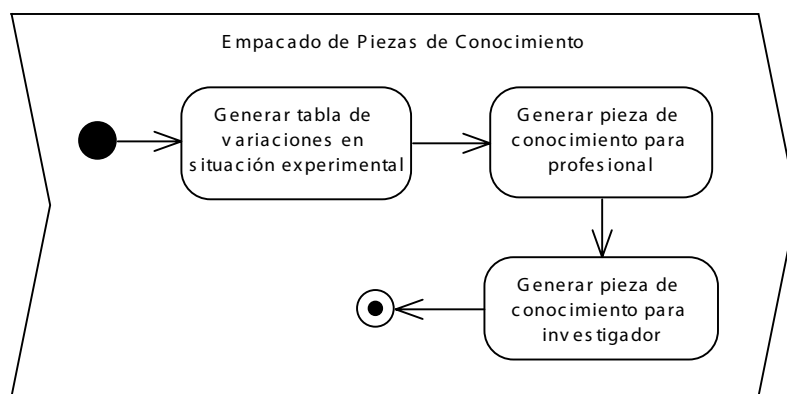


Fig. 9.17: Sub-proceso empaquetado de piezas de conocimiento

9.5.1 Generar Tabla de Variaciones en Situación Experimental

La primera actividad de este sub-proceso consiste en utilizar la información de las tablas de particiones generadas en el sub-proceso *identificación de similitudes* para calcular los cambios realizados en la situación experimental de las replicaciones que constituyen los

resultados consistentes. En cada sección de la situación experimental se identifica el número de cambios realizados en las replicaciones con resultados consistentes y el total de posibles cambios, de forma que se pueda calcular un porcentaje que indique el nivel de cambios realizados en cada elemento de la situación experimental y en cada sección. En esta actividad se utiliza el formulario PCR-10 ubicado en el apéndice D.10 para registrar los cambios así como los porcentajes de cambios realizados en los elementos y secciones de la situación experimental.

En la tabla 9.8 se muestra un ejemplo de la estructura principal del formulario PCR-10 con la información de los cambios y el porcentaje de variaciones realizadas en las replicaciones de la partición que tienen resultados consistentes.

Información de la partición					
Id Partición:	01				
Partición:	Efectividad/ Medida A/ Téc. A vs Téc. B/ T. Insp. A v1 vs. T. Insp. B v1				
Replicaciones consideradas para la pieza de conocimiento:	2 (r1, r2)				
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	1	1	100%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	1	1	100%	100%
Sitio		1	1	100%	100%
Protocolo experimental	Análisis	1	1	100%	
	Diseño	4	6	67%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	33%
Operacionalizaciones de los constructos	Tratamiento A	0	5	0%	
	Tratamiento B	0	5	0%	
	Efecto	0	2	0%	0%
Poblaciones	Sujetos	0	3	0%	
	Objetos experimentales	0	3	0%	0%

Tabla 9.8: Tabla de variaciones en situación experimental

La tabla 9.8 se divide en dos secciones, la primera sección contiene información de la partición seleccionada, las replicaciones usadas para empaquetar la pieza de conocimiento así como las replicaciones apartadas en caso de haber obtenido resultados inconsistentes. Esta información se obtiene a partir de la tabla resultante del análisis de resultados experimentales. La segunda sección de la tabla contiene información de las variaciones realizadas en la situación experimental. Por ejemplo, en la tabla 9.9 se muestra parte de la información de las tablas de particiones, en este caso las secciones de experimentadores y sitio.

En el caso del elemento *diseñador* no contiene sub-elementos, respecto a las replicaciones con resultados consistentes ($r1$ y $r2$) se observa que el elemento ha variado, como sólo hay un elemento, entonces se tiene un elemento variado de un total de un elemento i.e. se ha variado al 100% este elemento. De manera similar se calcula el resto de elementos. Por ejemplo, en la sección del *protocolo experimental* el elemento *diseño* se conforma de 6 sub-elementos, de acuerdo a las replicaciones seleccionadas para empaquetar la pieza

Estructura	Elemento Principal	ID	Categorías	r1	r2	r3
Experimentadores	Diseñador	1.1	Experimentador A	X		
		1.1	Experimentador C		X	X
	Entrenador	1.2	Experimentador A	X		
		1.2	Experimentador C		X	X
		1.2	Experimentador D			X
	Monitor	1.3	Experimentador A	X		
		1.3	Experimentador B	X		
		1.3	Experimentador C		X	X
		1.3	Experimentador D			X
	Medición	1.4	Experimentador B	X		
		1.4	Experimentador C		X	X
Sitio	Sitio	1.5	Experimentador B	X		
		1.5	Experimentador C		X	X
		2	Sitio A	X		
		2	Sitio B		X	X

Tabla 9.9: Tabla de partición referente a experimentadores y sitio

de conocimiento, se han variado cuatro de los seis sub-elementos por lo que se tiene en el elemento *diseño* una variación del 66.77%. Además de las variaciones por cada elemento de la situación experimental, también se calcula el porcentaje de variación de cada sección. En el ejemplo se observa que se ha variado al 100% los experimentadores y el sitio, y un 33% el protocolo experimental.

9.5.2 Generar Pieza de Conocimiento para Profesional

Una vez que se tiene la tabla con las variaciones de los elementos de la situación experimental es posible empaquetar las piezas de conocimiento. Las piezas de conocimiento se empaquetan con la información de la tabla de variaciones generada en la actividad anterior y con la tabla resultante del sub-proceso de *análisis de resultados experimentales*. En esta actividad se genera la pieza de conocimiento para el profesional. La pieza de conocimiento se compone de tres secciones que contienen información sobre la partición, los resultados de las comparaciones del par de tratamientos y de los cambios realizados en la situación experimental.

La primera sección identificada como *información de la partición* contiene los siguientes elementos:

- Id. Partición. Que representa el identificador de la partición.
- Partición. Que es una descripción de la partición utilizada en este sub-proceso.
- Replicaciones de la partición. Representa el número de replicaciones de la partición así como sus identificadores.
- Resultados de la identificación de similitudes. En este elemento se describe si en el sub-proceso *identificación de similitudes* se observaron cambios o no, en los elementos de la situación experimental.
- Tipo de resultados encontrados. Indica si se observaron o no inconsistencias en los resultados de la partición seleccionada.
- Tipo de inconsistencia observada. En este elemento se describen las características de la inconsistencia por ejemplo presencia de heterogeneidad debido a una de las replicaciones de la partición.

En la segunda sección *información de la comparación* se compone de los siguientes elementos:

- Replicaciones consideradas para la pieza de conocimiento. En este elemento se indican las replicaciones de la partición que se usarán para generar la pieza de

conocimiento. Se especifica el número de replicaciones seleccionadas y sus identificadores.

- Comparación. Indica si el efecto del tratamiento A es mayor, menor o igual que el tamaño del efecto del tratamiento B.
- Tipo de efecto global. Este elemento indica si el tamaño de efecto global estimado es o no significativo.
- Tamaño del efecto global. Indica el tamaño del efecto global estimado así como los intervalos de confianza estimados para el tamaño del efecto global verdadero.

En la tercera sección *variaciones en la situación experimental* se indica el nivel de variación por cada una de las secciones de la situación experimental y el tipo de evento o resultado experimental que se obtiene. El tipo de evento se obtiene de la tipología de replicaciones descrita en el capítulo 5. En esta sección se detalla el grado de verificación logrado en la comparación de replicaciones elegidas para empaquetar la pieza de conocimiento. Esta sección también contiene dos elementos que son el poder de confirmación de los resultados que indica el promedio de los porcentajes de variación de las secciones de la situación experimental. Entre mayor sea este promedio, la confirmación de los resultados es mayor, ya que el resultado observado es corroborado variando diferentes elementos de la situación experimental. El último elemento de esta sección indica la confiabilidad de la pieza de conocimiento, que es fiable si los resultados experimentales son consistentes.

Para construir la pieza de conocimiento se utiliza el formulario PCR-11 que se encuentra disponible en el apéndice D.11 de este documento. En la tabla 9.10 se muestra un ejemplo de la estructura principal del formulario PCR-11 con la pieza de conocimiento construida para el profesional.

De manera general la información que ofrece esta pieza de conocimiento al profesional es la siguiente. De tres replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al apartar la replicación $r3$ se obtienen resultados consistentes por lo que se incluyen sólo las replicaciones $r1$ y $r2$ para empaquetar la pieza de conocimiento. Los resultados de la comparación entre la técnica A y técnica B indican que la técnica A es significativamente más efectiva que la técnica B. Este resultado es objetivo independiente de los experimentadores y global independiente del sitio. Sin embargo el resultado observado puede no ser completamente real debido a que no se ha variado la mayor parte del protocolo experimental. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas para la pieza de conocimiento y el poder de confirmación de los resultados es del 47%.

9.5.3 Generar Pieza de Conocimiento para Investigador

La última actividad de este sub-proceso consiste en empaquetar la pieza de conocimiento para el investigador. Esta pieza de conocimiento es similar a la generada para el profesional, sin embargo, esta pieza contiene información de interés para el investigador como son los tipos de replicaciones realizados y en caso de haber resultados inconsistentes, información sobre el análisis de causas. La pieza de conocimiento para el investigador se compone de cuatro secciones que son: información de la partición, información de la comparación, variaciones en la situación experimental y análisis de causas.

La primera sección identificada como *información de la partición* contiene los siguientes elementos:

- Id. Partición. Que representa el identificador de la partición.

Información de la Partición		
Id Partición:	01	
Partición:	Efectividad/ Medida A/ Téc. A vs Téc. B/ T. Insp. A v1 vs. T. Insp. B v1	
Replicaciones de la partición:	3 (r1, r2, r3)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	Uno de los resultados es distinto del resto (situación 1)	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	2 (r1, r2)	
Comparación:	T. Insp. A es mayor que T. Insp. B	
Tipo de efecto global:	Siginiicativo	
Tamaño del efecto global:	0.3384 (pequeño) [0.0233; 0.6536]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	100%	Objetivo
Sitio:	100%	Global
Protocolo experimental:	33%	No es completamente real
Operacionalizaciones de los constructos:	0%	Limitado a la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	47%	
Confiabilidad de la pieza de conocimiento:	Confiable (Se ha excluido un experimento)	

Tabla 9.10: Pieza de conocimiento para profesional

- Partición. Que es una descripción de la partición utilizada en este sub-proceso.
- Replicaciones de la partición. Representa el número de replicaciones de la partición así como sus identificadores.
- Resultados de la identificación de similitudes. En este elemento se describe si en el sub-proceso *identificación de similitudes* se observaron cambios o no, en los elementos de la situación experimental.
- Tipo de resultados encontrados. Indica si se observaron o no inconsistencias en los resultados de la partición seleccionada.
- Información de los resultados. En este elemento se describen los resultados tanto de los gráficos de control como del meta-análisis. Respecto a los gráficos de control se especifican los límites de control superior e inferior, el promedio y las observaciones. En el meta-análisis se especifican los resultados de la prueba de heterogeneidad (estadístico Q y su valor p) así como los estadísticos I^2 y T^2 .
- Tipo de inconsistencia observada. En este elemento se describen las características de la inconsistencia por ejemplo presencia de heterogeneidad debido a una de las replicaciones de la partición.

En la segunda sección *información de la comparación* se compone se los siguientes elementos:

- Replicaciones consideradas para la pieza de conocimiento. En este elemento se indican las replicaciones de la partición que se usarán para generar la pieza de

conocimiento. Se especifica el número de replicaciones seleccionadas y sus identificadores.

- Comparación. Indica si el efecto del tratamiento A es mayor, menor o igual que el tamaño del efecto del tratamiento B.
- Tipo de efecto global. Este elemento indica si el tamaño de efecto global estimado es o no significativo.
- Tamaño del efecto global. Indica el tamaño del efecto global estimado así como los intervalos de confianza estimados para el tamaño del efecto global verdadero.

En la tercera sección *variaciones en la Situación Experimental* se indica el nivel de variación por cada una de las secciones de la situación experimental, el tipo de evento o resultado experimental que se obtiene de los diferentes cambios y los tipos de replicaciones identificados. En esta sección se detalla el grado de verificación logrado en la comparación de replicaciones elegidas para empaquetar la pieza de conocimiento, y al igual que en la pieza de conocimiento para el profesional, en esta sección se presenta el poder de confirmación de resultados y la confiabilidad de la pieza de conocimiento.

Por último, en la cuarta sección *Análisis de Causas* se describen las nuevas hipótesis de trabajo resultantes tras realizar el análisis de posibles causas que se cree influyen en los resultados de la partición. Para construir esta pieza de conocimiento se usa el formulario PCR-12 que se encuentra disponible en el apéndice D.12 de este documento. En la tabla 9.11 se muestra un ejemplo del contenido de la pieza de conocimiento para el investigador.

De acuerdo a la información de esta pieza de conocimiento, el investigador obtiene los siguientes resultados. De la partición seleccionada se observa heterogeneidad en el meta-análisis debido al tamaño de efecto estimado de la replicación $r3$. Al apartar esta replicación los resultados son consistentes entre sí y permiten generar la pieza de conocimiento. De acuerdo a la comparación de tratamientos, la técnica A es significativamente más efectiva que la técnica B, se observa un tamaño de efecto global pequeño a favor de la técnica A. Este resultado es objetivo independiente de los experimentadores y global independiente del sitio. Los tipos de replicaciones identificados son: *ajena* al ser distintos los investigadores de la replicación y el experimento base, *externa* al variarse el sitio, y *artificial* al variarse algunos de los elementos que conforman el protocolo experimental. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 47%. De acuerdo al análisis de causas se cree que la experiencia de los sujetos influye en los resultados de estos tratamientos. Al contar con la replicación $r3$ que es donde se cambiaron las poblaciones, se recomienda realizar una replicación lo más similar a $r3$ para verificar los resultados i.e. que el tipo de sujeto influye en el par de técnicas examinadas.

Información de la Partición			
Id Partición:	01		
Partición:	Efectividad/ Medida A/ Téc. A vs Téc. B/ T. Insp. A v1 vs. T. Insp. B v1		
Replicaciones de la partición:	3 (r1, r2, r3)		
Resultados del análisis de similitudes:	Con Cambios		
Tipo de resultados encontrados:	Inconsistentes		
Información de los resultados:	Q=6.1, p=0.0466, I²=67.4%, T²= 0.0976 Téc. A: 32.23% [22.12; 42.34] obs: 34.1, 29.7, 32.9 Téc. B: 29.4% [2.27; 56.53] obs: 28.1, 22.6, 37.5		
Tipo de inconsistencia observada:	Heterogeneidad en meta-análisis, r2 difiere de er1 y r1		
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	2 (r1, r2)		
Comparación:	T. Insp. A es mayor que T. Insp. B		
Tipo de efecto global:	Significativo		
Tamaño del efecto Global:	0.3384 (pequeño) [0.0233; 0.6536]		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	100%	Objetivo	Ajena
Sitio:	100%	Global	Externa
Protocolo experimental:	33%	No es completamente real	Artificial
Operacionalizaciones de los constructos:	0%	Limitado a la forma de aplicar los tratamientos y medir los resultados	-
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes	-
Poder de confirmación de resultados	47%		
Confiabilidad de la pieza de conocimiento:	Confiable (Se ha excluido un experimento)		
Análisis de Causas			
Nuevas hipótesis de trabajo:	La experiencia de los sujetos influye en los resultados de estas dos técnicas		

Tabla 9.11: Pieza de conocimiento para investigador

Parte III

Evaluación

Capítulo 10

Evaluación de Particiones

RESUMEN: A partir de un conjunto de replications en IS, en este capítulo comienza la evaluación de la solución propuesta.

En éste y en los sucesivos cinco capítulos se evalúa el marco conceptual para la comparación de replications en ISE así como la tipología de replications para la ISE. La evaluación se realiza siguiendo el procedimiento descrito en el capítulo 7 que usa las ideas clave de este marco conceptual. La evaluación se lleva a cabo utilizando un conjunto de replications en IS que a continuación se describe.

10.1 Conjunto de Replicaciones Estudiado

El conjunto de replications elegido para la evaluación se basa en el experimento realizado por Basili y Selby (Basili y Selby, 1987). El objetivo de este experimento es evaluar la efectividad, eficiencia y coste de tres técnicas de evaluación de software, específicamente prueba funcional (caja negra) utilizando particiones de clases de equivalencia PCE y análisis de valores límite AVL (Howden, 1980; Myers, 1979), prueba estructural (caja blanca) utilizando cobertura de sentencias (Marick, 1995; Myers, 1979) y lectura de código por abstracciones sucesivas (Linger et al., 1979)¹.

Basili y Selby (Basili y Selby, 1987) realizaron dos replications de este experimento. Después Kamsties y Lott (Kamsties y Lott, 1995b,a) llevaron a cabo dos replications. Más tarde, Roper et al. (Roper et al., 1997) realizaron otra replicación. Juristo y Vegas (Juristo y Vegas, 2003, 2009) también han realizado replications de este experimento.

No todos los experimentos que se han replicado toman como referencia el mismo experimento base. Por ejemplo, los experimentos de Kamsties y Lott (Kamsties y Lott, 1995b,a) replican los experimentos de Basili y Selby (Basili y Selby, 1987), el experimento de Roper et al. (Roper et al., 1997) replica los experimentos de Kamsties y Lott (Kamsties y Lott, 1995b,a).

Los experimentos de Juristo y Vegas (Juristo y Vegas, 2003, 2009) replican los experimentos de Kamsties y Lott (Kamsties y Lott, 1995b,a). Las replications en IS no parecen

¹En el apéndice E se describe de manera general el funcionamiento de estas técnicas.

cumplir la propiedad transitiva dados los cambios que suele introducir una replicación, no es lo mismo replicar un experimento que replicar la replicación de un experimento.

Dada la preeminencia del experimento original de Basili y Selby (Basili y Selby, 1987), así como la circunstancia de haberse replicado usando paquetes de laboratorio (y por tanto asumiendo que se producen pocos cambios en la replicación), algunas replicaciones se han usado por otros investigadores para agregar los resultados experimentales (Miller, 2000; Runeson et al., 2006). Sin embargo ambos intentos resultan fallidos. Los autores concluyen que los resultados de estas replicaciones son difíciles de agregar debido a los cambios introducidos en las diversas replicaciones. Estas circunstancias han llevado a elegir este conjunto de replicaciones como banco de pruebas para evaluar la tipología, el marco conceptual así como el procedimiento para la comparación de replicaciones propuestos en esta investigación.

El conjunto de replicaciones usado para evaluar el marco conceptual se conforma de 20 replicaciones donde cada una se identifica con la abreviación del lugar y el año en que se llevó a cabo. En la tabla 10.1 se listan las replicaciones elegidas².

Id Experimento	Lugar del experimento	País	Año
umd82	Universidad de Maryland	EE. UU.	1982
umd83	Universidad de Maryland	EE. UU.	1983
umd84	NASA	EE. UU.	1984
ukl94	Universidad de Kaiserslautern	Alemania	1994
ukl95	Universidad de Kaiserslautern	Alemania	1995
uos97	Universidad de Strathclyde	Reino Unido	1997
upm01	Universidad Politécnica de Madrid	España	2001
upm02	Universidad Politécnica de Madrid	España	2002
upm03	Universidad Politécnica de Madrid	España	2003
upm04	Universidad Politécnica de Madrid	España	2004
upm05	Universidad Politécnica de Madrid	España	2005
upv05	Universidad Politécnica de Valencia	España	2005
uds05	Universidad de Sevilla	España	2005
ort05	Universidad ORT	Uruguay	2005
upm06	Universidad Politécnica de Madrid	España	2006
upm07	Universidad Politécnica de Madrid	España	2007
upm08	Universidad Politécnica de Madrid	España	2008
upm09	Universidad Politécnica de Madrid	España	2009
upm10	Universidad Politécnica de Madrid	España	2010
upm11	Universidad Politécnica de Madrid	España	2011

Tabla 10.1: Replicaciones usadas para evaluar el marco conceptual para la comparación de replicaciones

La evaluación se realiza por partes y a través del procedimiento para comparar replicaciones. En este capítulo se evalúa el primer sub-proceso del procedimiento que es el *generador de particiones*. En el resto de capítulos de este apartado se evalúa el segundo sub-proceso *comparador de replicaciones*.

²En el apéndice F se encuentra disponible el material empleado en las replicaciones *upm06* a *upm10* que es similar al usado en el resto del conjunto de replicaciones.

10.2 Evaluación del Sub-proceso Generador de Particiones

Como se describió en el capítulo 7, el procedimiento para comparar replicaciones inicia con el sub-proceso *generador de particiones*. Para poder comparar varias replicaciones y obtener algún conocimiento de ellas, primero es necesario conocer las unidades comparables que tienen en común i.e. identificar las particiones. Las actividades que conforman este sub-proceso se listan a continuación:

1. Identificar constructos
2. Generar tabla de constructos
3. Identificar operacionalizaciones
4. Generar tabla de operacionalizaciones
5. Generar tabla de agrupaciones
6. Generar particiones
7. Seleccionar particiones

A continuación se llevan a cabo estas actividades en el conjunto de replicaciones antes descrito.

10.2.1 Identificar Constructos

En esta actividad se identifican los constructos de causa y efecto que son ideas o conceptos que en un experimento controlado se concretan a través de las operacionalizaciones. La identificación de constructos de causa y efecto se realizó en los artículos donde se describen este conjunto de replicaciones (Basili y Selby, 1987; Kamsties y Lott, 1995b,a; Roper et al., 1997; Juristo y Vegas, 2003, 2009). Los principales constructos de causa identificados son:

- Prueba funcional (Caja negra)
- Prueba estructural (Caja blanca)
- Lectura de código

Respecto a los constructos de efecto se identificaron:

- Efectividad
- Eficiencia
- Coste
- Visibilidad
- Aislamiento

10.2.2 Generar Tabla de Constructos

Una vez identificados los constructos se colocan en el formulario PCR-1 (apéndice D.1). En la tabla 10.2 se muestra la estructura principal de este formulario con la información de los constructos identificados en las replicaciones.

Identificador del experimento	Constructos de causa	Constructos de efecto
umd82..84	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
	Lectura de código	Coste
ukl94, ukl95	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
	Lectura de código	Aislamiento
uos97	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
	Lectura de código	
upm01..05	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
	Lectura de código	Visibilidad
ort05	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
upv05	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
		Visibilidad
uds	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
	Lectura de código	Visibilidad
upm06..11	Prueba funcional	Efectividad
	Prueba estructural	Eficiencia
	Lectura de código	

Tabla 10.2: Tabla de constructos del formulario PCR-1

10.2.3 Identificar Operacionalizaciones

En esta actividad se identifican las operacionalizaciones de los constructos que son los constructos concretizados i.e. los tratamientos como son aplicados por los sujetos. Las operacionalizaciones de causa identificadas en estas replicaciones son:

- Prueba funcional por PCE (Particiones de Clases de Equivalencia)
- Prueba funcional por PCE y AVL (Análisis de Valores Límite)
- Prueba estructural por cobertura de sentencias
- Prueba estructural por cobertura de predicados³
- Prueba estructural por cobertura de decisiones
- Lectura de código por abstracciones sucesivas

Respecto a las operacionalizaciones de efecto se identificaron:

- Faltas observadas, que son las faltas que el sujeto observa tras aplicar alguna de las pruebas

³Se ha abreviado como cobertura de predicados a esta técnica utilizada por (Kamsties y Lott, 1995b,a) donde de acuerdo a los autores se debe satisfacer el 100% la cobertura de: sentencias, condiciones múltiple, bucles y operadores relacionales.

- Faltas observables, que son las faltas observables en los casos de prueba⁴ especificados por el sujeto. Las faltas observables en un caso de prueba pueden no ser observadas por el sujeto
- Cobertura de sentencias, Que es el porcentaje máximo de sentencias de programa que fueron ejecutadas en un conjunto de casos de prueba
- Percepción de sujeto, que es la estimación de faltas que el sujeto cree haber detectado tras aplicar cada prueba
- Faltas observadas por tipo de falta
- Faltas observables por tipo de falta
- Tasa de detección de faltas, que son las faltas detectadas por el sujeto en un tiempo determinado
- Tasa de aislamiento, que son las faltas que el sujeto identificó en el código fuente suministrado en un tiempo determinado
- Casos de prueba generados, que es el número de casos de prueba que el sujeto generó durante la sesión del experimento
- Tiempo de detección de faltas, que es el tiempo que tardó el sujeto en aplicar cada técnica
- Faltas aisladas, que son las faltas que el sujeto identificó en el código fuente durante la sesión del experimento
- Faltas observadas/ faltas observables, que es una tasa que indica las faltas que fueron observadas por el sujeto del total de faltas observables definidas en los casos de prueba o en la especificación
- Faltas observadas con casos de prueba predefinidos, que son las faltas que el sujeto observó tras suministrarle un conjunto de casos de prueba predeterminado

Esta actividad puede implicar mayor esfuerzo si los autores de las replications emplean distintos términos para denotar un mismo concepto. En el caso de la operacionalización identificada como *faltas observadas* Kamsties y Lott (Kamsties y Lott, 1995b,a), Roper et al. (Roper et al., 1997) así como Juristo y Vegas (Juristo y Vegas, 2003, 2009) emplean esta operacionalización como *faltos observados*, mientras que Basili y Selby (Basili y Selby, 1987) la definen como *faltas detectadas*. Ya que las faltas descritas en estas replications causan fallos que pueden ser observados por los sujetos, se mantiene una relación entre el fallo observado y su respectiva falta que lo provoca. En este sentido se ha elegido denotar esta operacionalización como *faltas observadas* que el sujeto puede observar a través del fallo generado por el programa respecto a su especificación.

10.2.4 Generar Tabla de Operacionalizaciones

Una vez identificadas las operacionalizaciones se colocan en el formulario PCR-2 (apéndice D.2). En la tabla 10.3 se muestra la estructura principal de este formulario con la información de las operacionalizaciones identificadas en las replications.

⁴En el caso de la lectura de código por abstracciones sucesivas las faltas observables que se encuentran en la especificación que el sujeto escribió.

Constructos de Causa				Constructos de Efecto				
Id. Exp.	Prueba funcional	Prueba estructural	Lectura código	Efectividad	Eficiencia	Coste	Aislamiento	Visibilidad
umd82..84	P.C.E. y A.V.L.	Cob. Sent.	Lec. C. Abs. Suc.	Faltas observadas Faltas observables Cobertura de sent. Percepción del sujeto	Tasa de detección de faltas	Tiempo de detección de faltas		Faltas observadas / observables
ukl94, ukl95	P.C.E. y A.V.L.	Cob. Predicados	Lec. C. Abs. Suc.	Faltas observadas	Tasa de detección de faltas Tasa de aislamiento	Tiempo de detección de faltas	Faltas aisladas	
uos97	P.C.E. y A.V.L.	Cob. Decisiones	Lec. C. Abs. Suc.	Faltas observadas	Tasa de detección de faltas			
upm01..05	P.C.E.	Cob. Decisiones	Lec. C. Abs. Suc.	Faltas observadas Faltas observables Faltas observadas por tipo de falta Faltas observables por tipo de falta	Casos de prueba generados			Faltas observadas con CP predefinidos
ort05	P.C.E.	Cob. Decisiones		Faltas observables Faltas observables por tipo de falta	Casos de prueba generados			
upv05	P.C.E.	Cob. Decisiones		Faltas observadas Faltas observables Faltas observadas por tipo de falta Faltas observables por tipo de falta	Casos de prueba generados			Faltas observadas con CP predefinidos
uds05	P.C.E.	Cob. Decisiones	Lec. C. Abs. Suc.	Faltas observables Faltas observables por tipo de falta	Casos de prueba generados			Faltas observadas con CP predefinidos
upm06..10	P.C.E.	Cob. Decisiones	Lec. C. Abs. Suc.	Faltas observadas Faltas observables Faltas observadas por tipo de falta Faltas observables por tipo de falta	Casos de prueba generados			
upm11	P.C.E.	Cob. Decisiones		Faltas observadas	Casos de prueba generados			

Tabla 10.3: Tabla de operacionalizaciones del formulario PCR-2

10.2.5 Generar Tabla de Agrupaciones

Después de identificar y colocar en los respectivos formularios los constructos y las operacionalizaciones del conjunto de replications, en esta actividad se organizan las replications de acuerdo a sus unidades comparables, donde cada parte se conforma por un constructo y operacionalización de efecto y un par de constructos y operacionalizaciones de causa. En la tabla 10.4 se muestra la estructura principal del formulario PCR-3 (apéndice D.3) con un fragmento de la información de las unidades comparables en estas replications⁵.

Id. Exp.	Cons. Efecto	Oper. Efecto	Par constructos de causa	Par de operacionalizaciones de causa
umd82, umd83, umd84	Efectividad	Faltas observadas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
		Cobertura de sent.	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
		Percepción del sujeto	Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
ukl94, ukl95	Eficiencia	Tasa de detección de faltas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
	Coste	Tiempo de detección de faltas	Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
	Visibilidad	Faltas observadas/observables	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
	Eficiencia	Tasa de detección de faltas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
		Tasa de aislamiento	Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
	Coste	Tiempo de detección de faltas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
	Aislamiento	Faltas aisladas	Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.

Tabla 10.4: Tabla de agrupaciones del formulario PCR-3

10.2.6 Generar Particiones

Tomando como referencia la tabla de agrupaciones, en esta actividad se generan y representan de manera gráfica las particiones. En la tabla 10.5 se muestra un fragmento de la tabla con las particiones identificadas⁶, en total se identificaron 47 particiones.

En la figura 10.1 se muestran las particiones resultantes representadas de manera visual a través de mapas estructurados.

⁵En el apéndice G se encuentra la tabla completa.

⁶En el apéndice G se encuentra disponible la tabla con todas las particiones identificadas.

Partición	Constructos de efecto	Operacionalizaciones de efecto	Par de constructos de causa	Par de operacionalizaciones de causa	Id. Experimentos	Experimentos por partición
01	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..11, upv05	12
02	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias	umd82..84	3
03	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados	ukl94, ukl95, uos97	3
04	Efectividad	Faltas Observadas	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A. S.	upm01..10	10
05	Efectividad	Faltas Observadas	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs. Lec. C. por A.S.	udm82..84, ukl94, ukl95, uos97	6
06	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..10, uos97	11
07	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Sentencias vs. Lec. C. por A.S.	umd82..84	3
08	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Predicados vs. Lec. C. por A.S.	ukl94, ukl95	2
09	Efectividad	Faltas Observables	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..10, upv05, uds05, ort05	13
10	Efectividad	Faltas Observables	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias	umd82..84	3
11	Efectividad	Faltas Observables	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..09, uds05	10
12	Efectividad	Faltas Observables	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs Lec. C. por A.S.	umd82..84	3
13	Efectividad	Faltas Observables	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..09, uds05	10
14	Efectividad	Faltas Observables	Estruct. vs Lec. Cód.	Cob. Sentencias vs. Lec. C. por A.S.	umd82..84	3
15	Efectividad	Faltas observadas por tipo de falta	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..10, upv05	11
16	Efectividad	Faltas observadas por tipo de falta	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..10, upv05	10
17	Efectividad	Faltas observadas por tipo de falta	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..10	10
18	Efectividad	Faltas observables por tipo de falta	Func. vs Estruct.	P.C.E. vs Cob. Decisiones	upm01..09, upv05, uds05, ort05	12
19	Efectividad	Faltas observables por tipo de falta	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..09, uds05	10
20	Efectividad	Faltas observables por tipo de falta	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..09, uds05	10
21	Efectividad	Cobertura de sentencias	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias	umd82..84	3
22	Efectividad	Cobertura de sentencias	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs. Lec. C. por A.S.	umd82..84	3
23	Efectividad	Cobertura de sentencias	Estruct. vs Lec. Cód.	Cob. Sentencias vs. Lec. C. por A.S.	umd82..84	3

Tabla 10.5: Particiones identificadas en el conjunto de replicaciones

Efectividad											
Faltas Observadas						Faltas Observables por Tipo de Falta					
Func. vs Estruct.			Func. vs Lec. Cod.			Func. vs Estruct.		Estruct. vs Lec. Cod.		Func. vs Lec. Cod.	
P.C.E. vs Cob. Decisiones			P.C.E. y A.V.L. vs C			P.C.E. vs Cob. Decisiones		Cob. Decisiones vs Lec.		P.C.E. vs Lec. C. por A.	
upm01	upm04	upm07	umd82	umd83	umd84	ort05	uds05	upm01	uds05	upm01	
upm02	upm05	upm09				upm02	upm03	upm04	upm02	upm03	
upm03	upm06	upm10	P.C.E. y A.V. uki94	ukl95	P.C.E. uos97	upm05	upm07	upm08	upm04	upm05	
		upv05				upm06	upm09	upv05	upm06	upm08	
									upm07	upm09	
Estruct. vs Lec. Cod.			Cob. Decisiones vs Lec. C. por A.S.			Func. vs Estruct.		Estruct. vs Lec. Cod.		Func. vs Lec. Cod.	
Cob. Decisiones vs Lec. C. por A.S.			Cob. Sentencias v			P.C.E. vs Cob. Decisiones		Cob. Decisiones vs Lec.		P.C.E. y A.V. vs Lec. C. p	
uos97	upm03	upm06	umd82	umd84		upm01	upm03	upm05	upm07	upm09	upm10
upm01	upm04	upm08	umd83			upm02	upm04	upm06	upm08	upm10	upv05
upm02	upm05	upm09	Cob. Predicados v uki94	ukl95							
Faltas Observables						Faltas Observables por Tipo de Falta					
Func. vs Estruct.			Estruct. vs Lec. Cod.			Func. vs Estruct.		Estruct. vs Lec. Cod.		Func. vs Lec. Cod.	
P.C.E. vs Cob. Decisiones			Cob. Decisiones vs Lec. C.			P.C.E. vs Cob. Decisiones		Cob. Decisiones vs Lec.		P.C.E. y A.V. vs Lec. C. p	
ort05	uds05	upm01	uds05	upm01	upm02	upm01	upm03	upm05	upm07	upm09	upm10
upm02	upm05	upm06	upm03	upm06	upm07	upm02	upm04	upm06	upm08	upm10	upv05
upm03	upm07	upm09	upm04	upm08	upm09						
upm04	upm08	upv05	upm05	upm08	upm09						
P.C.E. y A.V.L. vs Cob. Sentencia	umd82	umd83	umd82	umd83	umd84						
Eficiencia						Eficiencia					
Tasa de Detección de Faltas			Casos de Prueba Generados			Tasa de Detección de Faltas		Casos de Prueba Generados		Func. vs Estruct.	
Estruct. vs Lec. Co			Func. vs Estruct.			Cob. Sentencias		P.C.E. y A.V.L. vs		P.C.E. vs Cob. Decisiones	
umd82	umd84		umd82	umd84		umd83		umd83		ort05	uds05
										upm02	upm03
										upm04	upm05
										upm06	upm07
										upm08	upm10
										upm09	upv05
Func. vs Lec. Cod.			Cob. Predicados vs			Func. vs Estruct.		Estruct. vs Lec. Cod.		Func. vs Lec. Cod.	
Cob. Predicados vs			P.C.E. y A.V.L. vs Co			P.C.E. y A.V.		Cob. Senten		P.C.E. y A.V.	
ukl94	ukl95		ukl94	ukl95		ukl94	ukl95	umd82	umd83	umd82	umd83
								umd84	umd84		
Coste						Visibilidad					
Tiempo de Detección de Faltas			Faltas Observadas con			Func. vs Estruct.		Func. vs L		Func. vs A	
Estruct. vs Lec. Cod.			Func. vs Estruct.			P.C.E. vs Cob. Decisi		P.C.E. y		P.C.E. y A	
Cob. Sentencias v	umd82	umd83	uds05	upm01	upm02	uds05	upm01	upm02	umd82	umd83	umd84
Cob. Predic	ukl94	ukl95	upm03	upm05	upm06	upm03	upm05	upm06	umd83	umd84	umd85
			upm04	upv05		upm04	upv05		umd84	umd85	umd86
Aislamiento						Aislamiento					
Faltas Aisladas			Func. vs Estruct.			Func. vs Estruct.		Func. vs L		Func. vs A	
Estruct. vs			P.C.E. y			P.C.E. y		P.C.E. y A		P.C.E. y A	
Cob. Pre	ukl94	ukl95	ukl94	ukl95		ukl94	ukl95	ukl94	ukl95	ukl94	ukl95

Fig. 10.1: Particiones resultantes del conjunto de replicaciones

En la figura 10.2 se muestran las replicasiones que emplean la eficiencia como constructo de efecto. En esta figura se han utilizado tres tonos de color para identificar las operacionalizaciones de efecto identificadas en estas replicasiones. Por ejemplo, la operacionalización de efecto identificada como *tasa de detección de faltas* contiene cinco particiones, cuatro particiones se componen de tres replicasiones cada una y la quinta partición se compone se seis replicasiones (*umd82..84*, *ukl94*, *ukl95* y *uos97*).

Eficiencia		Casos de Prueba Generados	
Tasa de Detección de Faltas		Func. vs Estruc.	
Estruc. vs Lec. Cod.	Cob. Sentencias vs Lec. C. por A.S.	P.C.E. y A.V.L. vs Cob. Sentencias	
	umd82	umd82	umd84
	umd83	umd83	
	Cob. Predicados vs Lec. C. Cob. Decisio	P.C.E. y A.V.L. vs Cob. Pre	P.C.E. y A.V.
		ukl94	uos97
		ukl95	
Func. vs Lec. Cod.		P.C.E. y A.V.L. vs Lec. C. por A.S.	
		umd82	umd84
		ukl94	
		ukl95	uos97
		umd83	
Tasa de Aislamiento de Faltas		Func. vs Estruc.	
Estruc. vs Lec. Cod.	Cob. Predicados vs Lec. C. po	P.C.E. y A.V.L. vs Cob. Predic	
	ukl94	ukl94	ukl95
	ukl95	ukl95	
Func. vs Lec. Cod.		P.C.E. y A.V.L. vs Lec. C. por	
		ukl94	ukl95
Casos de Prueba Generados		Func. vs Estruc.	
ort05	upm02	upm03	upm04
uds05	upm05	upm07	upm09
upm01	upm06	upm08	upm11
		upm10	upv05

Fig. 10.2: Particiones organizadas de acuerdo a las operacionalizaciones de efecto del constructo eficiencia

10.2.7 Seleccionar Particiones

Esta es la última actividad del *generador de particiones*. En esta actividad se seleccionan las particiones a utilizar en el sub-proceso *comparador de replications*. Como se describe en el capítulo 7 no todas las particiones se utilizan en el *comparador de replications*, sólo aquellas que contienen tres o más replications en la partición. A excepción de las particiones 08, 28, 29, 32, 33, 39, 42, 45, 46 y 47 el resto de particiones (que son 37) pueden usarse en el *comparador de replications*. De las 37 particiones disponibles, de manera inicial se han seleccionado para la evaluación cuatro particiones (01, 06, 09, 11) más otra partición (la 48) generada a partir de los constructos con el fin de evaluar una posible generalización de resultados a nivel de operacionalizaciones. Estas particiones contienen diferentes tipos de replications identificadas con la tipología propuesta como: replications literales-nativas-internas, diferenciadas-nativas-internas, diferenciadas-conjuntas-externas así como replications diferenciadas-ajenas-externas.

En la tabla 10.6 se muestran las particiones inicialmente seleccionadas para evaluar el sub-proceso *comparador de replications*. La partición 48 se ha construido a partir de las particiones 01, 02 y 03. Esta partición se genera a partir de los constructos y no de las operacionalizaciones como sucede con el resto de particiones.

Partición	Constructos de efecto	Operacionalizaciones de efecto	Par de constructos de causa	Par de operacionalizaciones de causa	Id. Experimentos	Experimentos por partición
01	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..11, upv05	12
06	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..10, uos97	11
09	Efectividad	Faltas Observables	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..10, upv05, uds05, ort05	13
11	Efectividad	Faltas Observables	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..09, uds05	10
48	Efectividad	Faltas Observadas	Func. vs Estruct.	-	upm01..11, upv05, umd82..84, ukl94, ukl95, uos97	18

Tabla 10.6: Conjunto inicial de particiones seleccionadas

A partir del conjunto inicial de particiones seleccionadas, se han dividido algunas de las particiones para evaluar el *comparador de replications* bajo distintas situaciones. En la tabla 10.7 se muestran las cinco particiones usadas para evaluar el *comparador de replications*.

Partición	Constructos de efecto	Operacionalizaciones de efecto	Par de constructos de causa	Par de operacionalizaciones de causa	Id. Experimentos	Experimentos por partición
01.a	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..05	5
06.b	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..10	10
11.c	Efectividad	Faltas Observables	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..09	9
09	Efectividad	Faltas Observables	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..10, upv05, uds05, ort05	13
48	Efectividad	Faltas Observadas	Func. vs Estruct.	-	upm01..11, upv05, umd82..84, ukl94, ukl95, uos97	18

Tabla 10.7: Particiones seleccionadas para usar en el *comparador de replications*

Cada una de las particiones representa una situación particular para evaluar. En la primera situación se usa la partición 01.a para evaluar un conjunto de replications que se han realizado de manera literal en el mismo sitio sin variar el protocolo experimental, operacionalizaciones y poblaciones. En la segunda y tercera situación se utilizan las

particiones 06.b y 11.c para evaluar un conjunto de replicasiones diferenciadas que se han realizado en el mismo sitio, no obstante, se han variado elementos del protocolo experimental así como elementos de las poblaciones y con distintas métricas. En la cuarta situación se utiliza la partición 09 para evaluar un conjunto de replicasiones diferenciadas que se se han realizado de manera coordinada en diferentes sitios. Por último en la quinta situación se utiliza la partición 48 para evaluar un conjunto de replicasiones diferenciadas que se han realizado en distintos sitios por distintos experimentadores con variaciones en el protocolo experimental, operacionalizaciones y poblaciones.

La evaluación del *comparador de replicasiones* con las particiones seleccionadas se describe en los siguientes capítulos. En cada capítulo se desarrolla la evaluación del sub-proceso con una de las cinco particiones seleccionadas.

10.3 Resultados de la Evaluación

El *generador de particiones* ha permitido agrupar en particiones las unidades comparables comunes entre las replicasiones seleccionadas. La representación visual de las particiones permite al investigador conocer los distintos tratamientos y métricas (constructos y operacionalizaciones de causa y efecto) usados en el conjunto de replicasiones. Las particiones resultantes de este sub-proceso permiten identificar las diferentes comparaciones que se pueden realizar entre los constructos y operacionalizaciones de causa y efecto identificados.

Nótese que el estado de la práctica en la generación de conocimiento a partir de replicasiones era que se intentaba combinar replicasiones completas. Dada la gran cantidad de variaciones que puedan contener las replicasiones, y que aún así contengan partes equivalentes se hacia inviable la comparación de resultados. Este sub-proceso permite al investigador gestionar una gran cantidad de información guiándole a través de ella hasta lograr *separar la paja del heno* quedándose con información relevante en este punto del procedimiento.

Capítulo 11

Evaluación con Replicaciones Literales

RESUMEN: En este capítulo se evalúan las ideas del marco para la comparación de replicaciones referentes a las comparaciones entre replicaciones. Esta evaluación se realiza a través del sub-proceso *comparador de replicaciones* usando una partición compuesta por una serie de replicaciones realizadas de forma literal, nativa e interna.

Una vez que se dispone de una o varias particiones, el siguiente sub-proceso del procedimiento para la comparación de replicaciones orienta al investigador en la comparación de situaciones y resultados experimentales, y en base al grado de verificación alcanzado se obtiene conocimiento útil para el profesional y el investigador.

En este capítulo se evalúa el *comparador de replicaciones* con la partición 01.a¹. Esta partición se compone de cinco replicaciones literales, nativas e internas (*upm01..05*) realizadas en cinco años consecutivos por el grupo de investigación GrISE (Grupo de Investigación en Ingeniería del Software Empírica) de la FI-UPM (Facultad de Informática de la Universidad Politécnica de Madrid). En esta partición se compara la efectividad medida como el porcentaje de faltas observadas de las pruebas PCE y cobertura de decisiones.

El *comparador de replicaciones* se compone de la actividad *generar tabla maestra* y de los sub-procesos: *identificación de similitudes*, *análisis de resultados experimentales* y *empaquetado de piezas de conocimiento*.

11.1 Generar Tabla Maestra

En esta actividad se extrae la información de los elementos de la situación experimental de las replicaciones que componen las particiones seleccionadas. Las particiones a utilizar en este sub-proceso son cinco (01.a, 06.b, 11.c, 09 y 48). En total son 20 las replicaciones que conforman las distintas particiones (*umd82..84*, *ukl94*, *ukl95*, *uos97*, *upm01..11*, *upv05*,

¹Con el fin de mantener una misma estructura y facilitar la redacción, en los cuatro capítulos sucesivos se usa como referencia la estructura de este capítulo.

uds05 y *ort05*). La información referente a las replications *umd82..84*, *ukl94*, *ukl95*, *uos97*, *upm01..05*, *upv05* y *uds05* se ha extraído de (Basili y Selby, 1987; Kamsties y Lott, 1995b,a; Roper et al., 1997; Juristo y Vegas, 2003, 2009). Respecto a las replications *upm06..11* la información se ha extraído de repositorios de datos que tiene el grupo de investigación GrISE de estas replications.

En la tabla 11.1 se muestra un fragmento de la tabla maestra generada referente a los elementos experimentadores y sitio. La información completa de esta tabla se encuentra disponible en el apéndice G. Esta tabla se compone de las tablas G.5, G.6, G.7, G.8, G.9, G.10, G.11, G.12, G.13, G.14, G.15 que usan como referencia el formulario PCR-5² (apéndice D.5) con la información extraída de estas replications. En estas tablas la información se ha colocado de acuerdo a los diferentes elementos de la situación experimental descritos en la sección 4.4 del capítulo 4.

Por limitaciones en las dimensiones del papel usado en este documento, algunas de las categorías de las tablas se han abreviado o contienen descripciones cortas. A continuación se describen algunas de estas categorías. En las tablas G.7, G.8, G.9 se describen las categorías referentes al protocolo experimental:

- Orden de aplicación de los tratamientos. La categoría *todos los tipos de orden* se refiere a las secuencias de asignación de los tratamientos (prueba funcional, estructural, lectura) a los sujetos. *Todos los tipos de orden* se refiere a las secuencias: Funcional (F) - Estructural (E) - Lectura (L), F-L-E, E-F-L, E-L-F, L-F-S y L-E-F. Estos ordenes de aplicación se asignan a los sujetos en las sesiones del experimento. En algunas replications como *upm11* sólo se tiene un orden de aplicación del tratamiento que es E-F, esto indica que los sujetos primero se les asignó la prueba estructural y después la funcional. En algunas replications no se tiene orden de aplicación porque los sujetos sólo se les asignó un tratamiento i.e. una de las técnicas de evaluación de software.
- Combinación del entrenamiento con las sesiones. Las categorías que se identificaron son *secuencial* e *intercalado*. *Secuencial* se refiere a que a los sujetos se les dio el entrenamiento y después se realizaron las sesiones del experimento, mientras que *intercalado* se refiere a que a los sujetos se les entrenó en alguno de los tipos de prueba y después se realizó la sesión del experimento referente al entrenamiento dado.
- Faltas (instancias). En los objetos experimentales (códigos fuente) se observaron diferentes faltas distribuidas de manera diferente. Estas instancias de faltas se clasificaron con letras donde cada letra indica una distribución de instancias diferente.
- Formularios del experimento. Respecto a las categorías de este elemento, se identificaron modificaciones en los formularios de recogida de datos como pueden ser inclusión o remoción de algún campo entre un experimento y otro. Estos cambios se han agrupado en versiones i.e. formularios v1 o formularios v2, etc.
- Instrucciones del experimento. De manera similar a los formularios, se observaron cambios en las instrucciones asignadas a los sujetos para que apliquen los tratamientos en este caso las técnicas de evaluación de software, al igual que los formularios, los cambios en las instrucciones se agruparon en versiones.

En las tablas G.10, G.11, G.12 se describen las categorías referentes a las operaciona-

²La información de cada replicación se ha pasado directamente a la tabla maestra por lo que se ha prescindido del formulario PCR-4 (apéndice D.4).

	Experimentadores y Sitio									
Elemento Principal	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04
Diseñador	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Vegas	Vegas	Vegas	Vegas
Entrenador	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Juristo y Vegas	Juristo y Vegas	Juristo y Vegas	Juristo y Vegas
Monitor	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Juristo y Vegas	Juristo, Vegas, Dieste y Ferré	Juristo y Vegas	Juristo y Vegas
Medición	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Vegas	Vegas	Vegas	Vegas
Analista	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Vegas	Vegas	Vegas	Vegas
Sitio	UMD	UMD	NASA	UKL	UKL	UoS	UPM	UPM	UPM	UPM

Tabla 11.1: Elementos identificados en experimentadores y sitio

lizaciones de los constructos:

- Tratamiento acorde a la literatura. Las categorías abreviadas en este elemento son:
 - PCE y AVL, que es, Particiones de Clases de Equivalencia y Análisis de Valores Límite
 - PCE, que es, Particiones de Clases de Equivalencia
- Procedimiento de aplicación del tratamiento. Respecto a las categorías de este elemento, se observaron cambios en la forma como el sujeto aplica los tratamientos, particularmente en la prueba funcional y estructural, por ejemplo, en la prueba funcional en algunas replicaciones los sujetos utilizan PCE y AVL mientras que en otras replicaciones sólo se usa PCE, de manera similar sucede con la prueba estructural. A estos cambios se identificaron como distintas versiones.
- Tratamiento como instrucciones. Al igual que en el elemento anterior, las categorías de este elemento se diferenciaron con versiones.

En las tablas G.13, G.14, G.15 se describen las categorías referentes a las poblaciones:

- Tipos de faltas. Estas categorías se refieren a las características de las faltas que el investigador usó en los objetos experimentales. Se observaron diferentes agrupaciones de faltas por lo que se utilizaron varias letras para su diferenciación.

Una vez finalizada esta actividad a continuación se evalúan los tres sub-procesos restantes (*identificación de similitudes, análisis de resultados experimentales y empaquetado de piezas de conocimiento*) con la partición 01.a.

11.2 Sub-proceso Identificación de Similitudes

Una vez que se ha rellenado la información de las particiones en la tabla maestra, el primer sub-proceso del *comparador de replicaciones* se usa para conocer los cambios y representar la similitud que hay entre las replicaciones de las particiones seleccionadas. Esta partición se compone de las replicaciones *upm01..05* por lo que sólo se usa la información de la tabla maestra referente a las replicaciones de esta partición. Este sub-proceso se conforma de las siguientes actividades:

- Generar tablas de particiones
- Identificar similitudes

11.2.1 Generar Tablas de Particiones

Las tablas de particiones sirven como referencia para generar las representaciones de similitudes que hay entre las situaciones experimentales de las replicaciones contenidas en una partición, en este caso, la partición 01.a. En la tabla 11.2 se muestra la tabla de partición referente a los experimentadores, sitio y protocolo experimental. Las tablas completas de esta partición se encuentran en el apéndice G.

11.2.2 Identificar Similitudes

En esta actividad del sub-proceso *identificación de similitudes* se usan las tablas de particiones más el procedimiento de automatización de cálculos para la identificación de similitudes descrito en el apéndice B. Este procedimiento facilita la generación de gráficos que representan las similitudes entre las replicaciones de una partición.

Sección	Elemento principal	Sub elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05
Experimentadores	Diseñador		1.1	Vegas	X	X	X	X	X
	Entrenador		1.2	Juristo	X	X	X	X	X
			1.2	Vegas	X	X	X	X	
			1.2	Moreno					X
	Monitor		1.3	Juristo	X	X	X	X	
			1.3	Vegas	X	X	X	X	X
			1.3	Dieste		X			
			1.3	Ferre		X			
			1.3	Moreno					X
	Medición		1.4	Vegas	X	X	X	X	
			1.4	Solari					X
	Analista		1.5	Vegas	X	X	X	X	X
Sitio			2	UPM	X	X	X	X	X
Protocolo experimental	Análisis	Técnica de análisis de datos	3.1.1	ANOVA factorial	X	X	X	X	X
	Diseño	Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X	X	X
		Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X	X	X
		Combinación del entrenamiento con las sesiones	3.2.3	Secuencial	X	X	X	X	X
		Tratamientos por sesión	3.2.4	3 Técnicas, 1 Programa	X	X	X	X	X
		Sesiones requeridas por sujeto para aplicar tratamientos	3.2.5	1 Sesión	X	X	X	X	X
		Tareas por sujeto por sesión	3.2.6	1 Técnica, 1 Programa	X	X	X	X	X
	Instrumentos	Instrumentos o equipo de medición	3.3.1	Ordenador, folios, bolígrafos (manual)	X	X	X	X	X
	Objetos experimentales	Programas	3.4.1	Cmdline	X	X	X	X	X
			3.4.1	Nametbl	X	X	X	X	X
			3.4.1	Ntree	X	X	X	X	X
		Faltas (Instancias)	3.4.2	Instancia G	X	X	X	X	X
			3.4.2	Instancia H	X	X	X	X	X
	Materiales	Formularios del experimento	3.5.1	Formularios v2	X	X	X	X	X
		Instrucciones del experimento	3.5.2	Instrucciones v2	X	X	X	X	X
		Guías suplementarias	3.5.3	Guías v2	X	X	X	X	X

Tabla 11.2: Tabla de particiones referente a experimentadores, sitio y protocolo experimental

Siguiendo la automatización de cálculos, primeramente se organizan las tablas de particiones de las diferentes secciones de la situación experimental en un fichero de texto CSV (*Comma-separated values*)³ que representa la partición 01.a.

El fichero CSV de la partición 01.a se ingresa como parámetro en el programa *Transpose* para generar ficheros de texto que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición. Los ficheros de texto resultantes se utilizan para generar de forma automática las representaciones de las similitudes. Este paso se realiza tras cargar el fichero `r_script01.txt` en el entorno R (R Dev. Core Team, 2011). Este fichero contiene varias funciones en R (R Dev. Core Team, 2011) para calcular las diferentes matrices de distancias así como las funciones para realizar los análisis de conglomerados jerárquico y de escalado multidimensional.

Una vez cargado en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`, se obtienen dos ficheros EPS (*Encapsulated PostScript*) que representan de forma gráfica las similitudes entre las replicaciones de la partición. En la figura 11.1 se muestran los gráficos resultantes que representan las similitudes de situaciones experimentales en esta partición⁴.

11.3 Sub-proceso Análisis de Resultados Experimentales

En este sub-proceso se comparan los resultados experimentales de las replicaciones de la partición y junto con los cambios identificados en las situaciones experimentales se determina el grado de verificación logrado en una partición. El sub-proceso *análisis de resultados experimentales* se compone de las siguientes actividades:

- Generar tabla de resultados experimentales
- Realizar comparación de resultados experimentales
- Generar tabla con información del análisis realizado

11.3.1 Generar Tabla de Resultados Experimentales

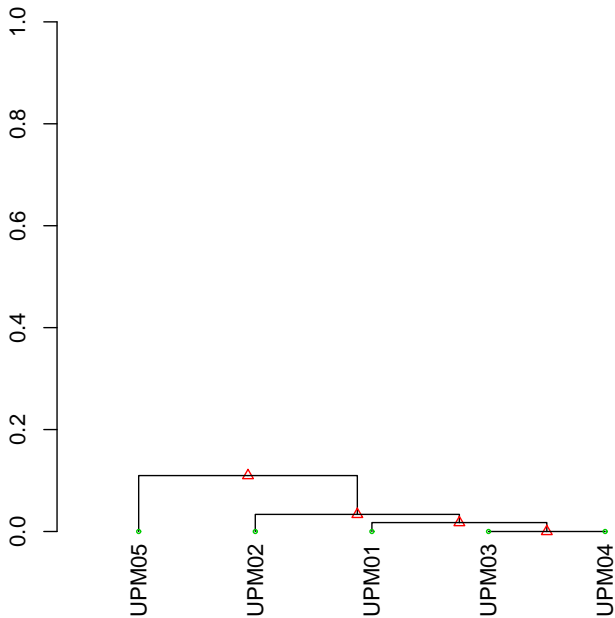
En esta actividad se extraen los resultados experimentales de las replicaciones que conforman la partición y se registran en el formulario PCR-7 descrito en el apéndice D.7. La tabla 11.3 muestra los resultados experimentales obtenidos de la partición 01.a.

Id exp.	P.C.E.			Cob. Decisiones		
	Tamaño (n_1)	Porcentaje faltas observadas ($\bar{\gamma}_1$)	Desv. est. (s_1)	Tamaño (n_2)	Porcentaje faltas observadas ($\bar{\gamma}_2$)	Desv. est. (s_2)
upm01	42	62.4	22.88	42	57.59	19.56
upm02	39	81.45	17.7	39	84.88	19.56
upm03	29	75.78	13.39	29	70.27	19.19
upm04	36	66.6	17.72	36	63.38	18.64
upm05	32	62.2	23.08	32	63.39	23.5

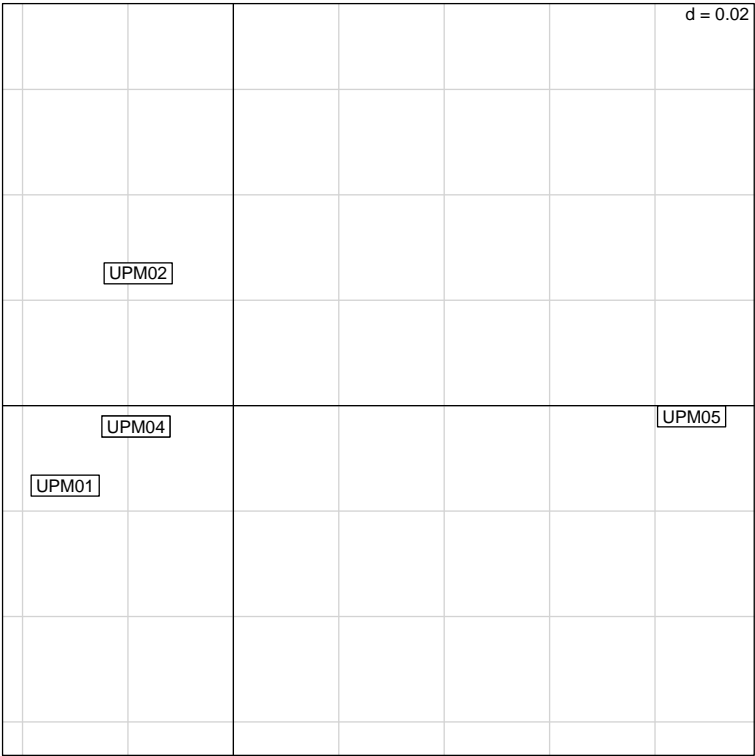
Tabla 11.3: Resultados experimentales de la partición 01.a

³En el primer enlace de la sección *resources* de la url <http://www.osgg.net> se encuentra disponible el fichero de texto CSV de esta partición, así como los ficheros generados de forma automática que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición.

⁴Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`.



(a) Dendrograma



(b) MDS

Fig. 11.1: Representación de similitudes, partición 01.a

11.3.2 Realizar Comparación de Resultados Experimentales

La siguiente actividad de este sub-proceso consiste en comparar los resultados experimentales de la partición. El análisis de los resultados se lleva a cabo a través del uso de gráficos de control y de técnicas estadísticas de agregación de resultados i.e. meta-análisis.

A partir de la tabla de resultados experimentales se generan dos gráficos de control, uno por cada tratamiento así como el meta-análisis que compara los resultados de los tratamientos mediante los tamaños de efecto estimados de las replicaciones de una partición así como el tamaño de efecto global estimado de la partición. Para estos cálculos se usa el procedimiento de automatización referente al análisis de resultados experimentales descrito en el apéndice B.

Estos cálculos se llevan a cabo tras ejecutar el fichero `r_script02.txt` en el entorno R (R Dev. Core Team, 2011) para la partición 01.a. Este fichero recibe como entrada la tabla de resultados experimentales de la partición. Como salida se generan tres ficheros EPS, dos ficheros que contienen los gráficos de control del par de tratamientos (PCE y cob. decisiones), y otro fichero que contiene los resultados del meta-análisis. Con los gráficos de control y meta-análisis resultantes más las representaciones de similitudes de situaciones experimentales obtenidas en actividades previas se está en situación de decidir el grado de verificación que alcanzan los resultados experimentales.

En la figura 11.2 se muestran los gráficos de control de los dos tratamientos examinados en esta partición (PCE y cob. decisiones), mientras que en la figura 11.3 se muestran los resultados del meta-análisis⁵.

Como se muestra en las figuras 11.2a y 11.2b, los resultados experimentales obtenidos en las cinco replicaciones para ambos tratamientos se encuentran dentro de los límites de control. Lo que significa que en principio los resultados son verificados en esta partición.

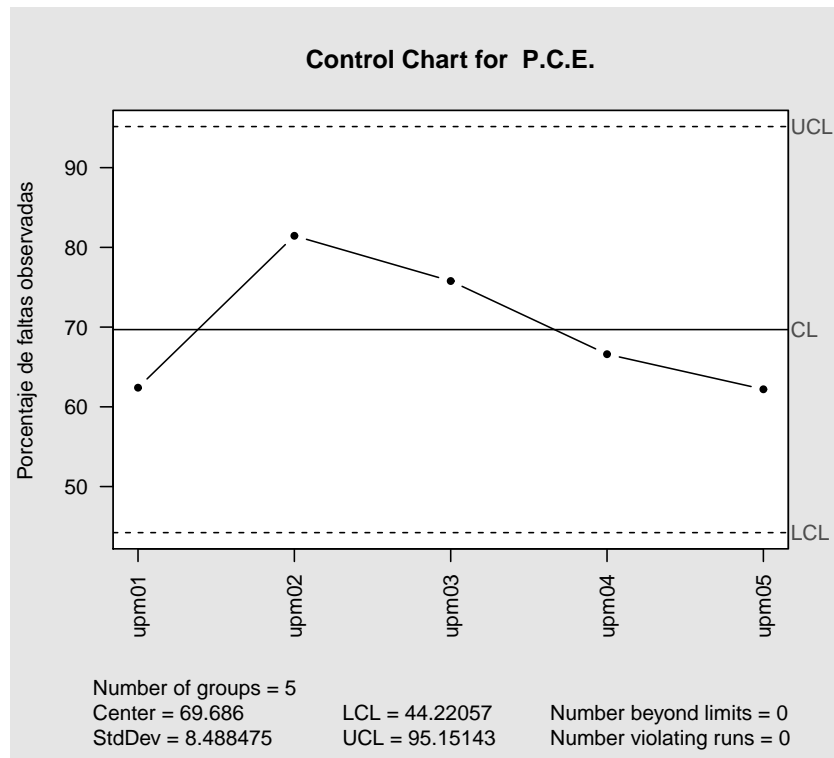
Los resultados del meta-análisis (figura 11.3) indican homogeneidad en la distribución de tamaños de efecto estimados. Es decir, no se observa heterogeneidad de resultados y por tanto se corrobora que los resultados pueden considerarse verificados⁶.

Aunque el efecto global estimado sugiere que la efectividad de PCE es ligeramente mejor que la efectividad de la prueba cobertura de decisiones, el valor cercano a cero (0.0924) del efecto global estimado indica que no hay diferencia significativa entre ambos tratamientos. De acuerdo a los intervalos de confianza estimados, el tamaño de efecto verdadero global puede ser cero. Por tanto, la efectividad de ambos tratamientos es la misma. Al estar las observaciones de los gráficos de control dentro de sus límites y al no observar heterogeneidad en los tamaños de efecto estimados en las replicaciones de esta partición, ambos resultados (gráficos de control y meta-análisis) se consideran consistentes.

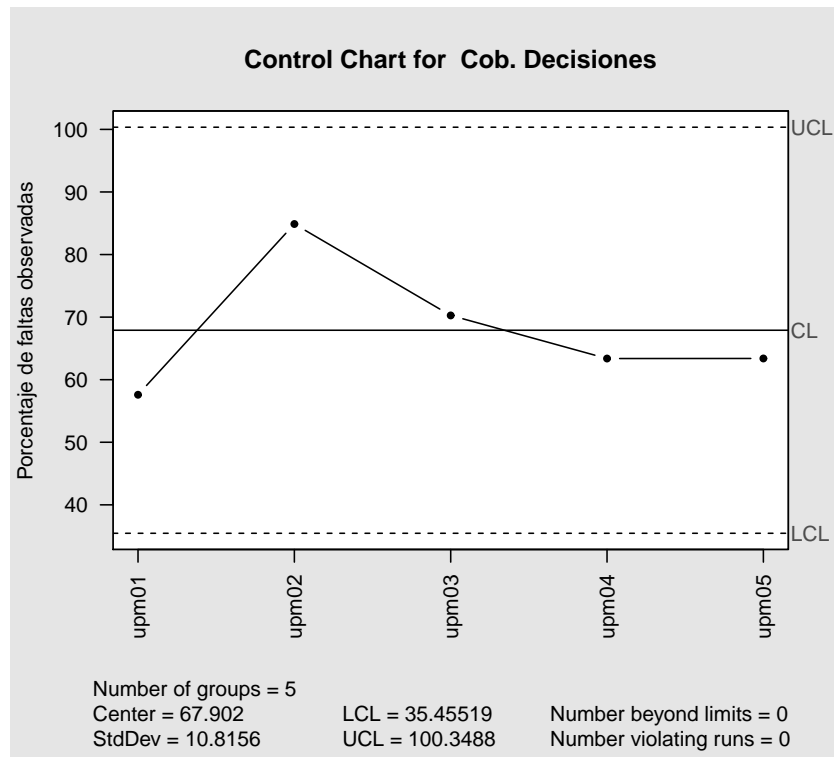
En el caso de los resultados del sub-proceso *identificación de similitudes* el dendrograma resultante de la figura 11.1a indica que las replicaciones de la partición son muy similares entre sí. Esto es que las replicaciones comparten en común la mayoría de los elementos de la situación experimental. La replicación *upm05* es ligeramente menos similar del resto, sin embargo en la escala de 0 a 1, donde 0 indica una total similitud y 1 una total dis-similitud, la replicación *upm05* se encuentra a una distancia menor a 0.2 del resto de replicaciones de la partición. En la figura 11.1b se muestra el MDS con otra

⁵Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script02.txt`.

⁶la prueba de heterogeneidad no es significativa $Q = 3.1$, $p = 0.5348$ de manera similar los resultados de los índices I^2 y T^2 ambos son igual a cero.



(a) Prueba por PCE



(b) Prueba por cob. de decisiones

Fig. 11.2: Gráficos de control resultantes

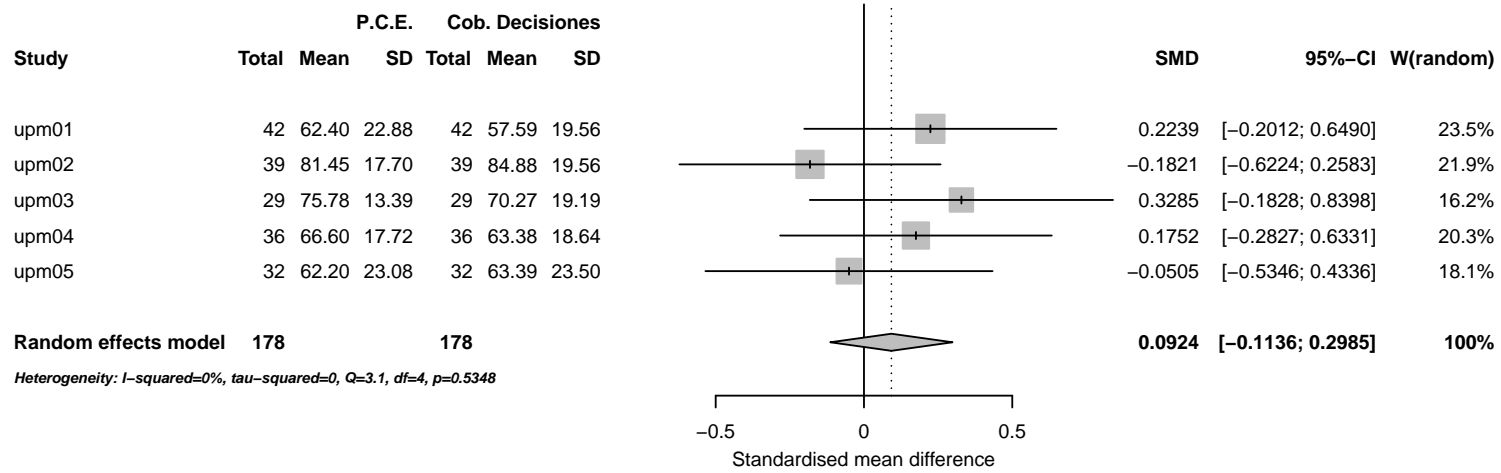


Fig. 11.3: Meta-análisis de la prueba por PCE y cob. de decisiones

representación de las similitudes. En esta figura al ser iguales las replicaciones *upm03* y *upm04*, la replicación *upm03* no es visible ya que se encuentra detrás de *upm04*.

De acuerdo a los resultados de los gráficos control y meta-análisis, se tienen resultados consistentes mientras que en la identificación de similitudes se tienen algunos cambios en los elementos de la situación experimental. El conocimiento que se obtiene es confiable ya que se tiene consistencia en los resultados además de obtener cierta ganancia de conocimiento al haber algunos cambios en los elementos de la situación experimental.

11.3.3 Generar Tabla con Información del Análisis Realizado

Al no encontrarse inconsistencias en los gráficos de control y el meta-análisis la siguiente actividad de este sub-proceso consiste en colocar en el formulario PCR-9 descrito en el apéndice D.9 la información de las comparaciones de resultados realizadas en la actividad anterior, esta información se utiliza para determinar el grado de verificación alcanzado y empaquetar este conocimiento. En la tabla 11.4 se muestra la estructura principal del formulario PCR-9 con la información de los resultados de las comparaciones realizadas en la actividad anterior.

Información de la Partición	
Id. Partición:	01.a
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.
Replicaciones de la partición:	5 (upm01..upm05)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Consistentes
Información de los resultados:	Q=3.1, p=0.5348, I ² =0%, T ² = 0 P.C.E.: 69.69% [44.22; 95.15] f. observadas: 62.4, 81.45, 75.78, 66.6, 62.2 Cob. Dec.: 67.9% [35.45; 100] f. observadas: 57.59, 84.88, 70.27, 63.38, 63.39
Tipo de inconsistencia observada:	inconsistencias no observadas
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	5 (upm01..upm05)
Comparación del par de tratamientos:	P.C.E. = Cob. Dec. (Sin diferencias en efectividad)
Tipo de efecto global:	No significativo
Tamaño del efecto global:	0.0924 [-0.1136; 0.2985]

Tabla 11.4: Información de las comparaciones de resultados realizadas en las replicaciones de la partición 01.a

11.4 Sub-proceso Empaquetado de Piezas de Conocimiento

Este es el último sub-proceso del *comparador de replicaciones* responsable de empaquetar las piezas de conocimiento. Este sub-proceso se compone de las siguientes actividades:

- Generar tabla de variaciones en situación experimental
- Generar pieza de conocimiento para profesional
- Generar pieza de conocimiento para investigador

11.4.1 Generar Tabla de Variaciones en Situación Experimental

En esta actividad se usa la información de las tablas de particiones generadas en el sub-proceso *identificación de similitudes* para calcular los cambios realizados en la situación experimental de las replicaciones que constituyen los resultados consistentes, en este caso las replicaciones *upm01..05* de la partición 01.a. En esta actividad se utiliza el formulario PCR-10 (apéndice D.10) para colocar los cálculos de los cambios realizados. La tabla resultante se utiliza en las siguientes actividades del sub-proceso para empaquetar las piezas de conocimiento. En la tabla 11.5 se muestra la estructura principal del formulario PCR-10 con la el cálculo de variaciones en las replicaciones de la partición 01.a.

Información de la partición					
Id Partición:	01.a				
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.				
Replicaciones consideradas para la pieza de conocimiento:	5 (upm01..upm05)				
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	0	1	0%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	0	1	0%	60%
Sitio		0	1	0%	0%
Protocolo experimental	Análisis	0	1	0%	
	Diseño	0	6	0%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	0%
Operacionalizaciones de los constructos	Tratamiento A	1	5	20%	
	Tratamiento B	1	5	20%	
	Efecto	0	2	0%	13%
Poblaciones	Sujetos	0	3	0%	
	Objetos experimentales	0	4	0%	0%

Tabla 11.5: Tabla de variaciones en situación experimental de las replicaciones upm01..05

11.4.2 Generar Pieza de Conocimiento para Profesional

A partir de la tabla con información del análisis realizado (sección 11.3.3) y con la tabla de variaciones (sección 11.4.1) se determina el grado de verificación logrado en esa partición y se empaqueta este conocimiento para que sea usado por el profesional. Para empaquetar

Información de la Partición		
Id Partición:	01.a	
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.	
Replicaciones de la partición:	5 (upm01..upm05)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Consistentes	
Tipo de inconsistencia observada:	no se observaron inconsistencias	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	5 (upm01..upm05)	
Comparación:	P.C.E. = Cob. Dec. (Sin diferencias en efectividad)	
Tipo de efecto global:	No Significativo	
Tamaño del efecto global:	0.0924 [-0.1136; 0.2985]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	60%	Semi-objetivo
Sitio:	0%	Local
Protocolo experimental:	0%	Artificial
Operacionalizaciones de los constructos:	13%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	15%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 11.6: Pieza de conocimiento para profesional

este conocimiento se usa el formulario PCR-11 (apéndice D.11). En la tabla 11.6 se muestra la estructura principal del formulario PCR-11 con la información de la pieza de conocimiento resultante para el profesional.

De manera general la información que ofrece esta pieza de conocimiento al profesional es la siguiente. De cinco replicaciones que conforman la partición seleccionada, se observan resultados consistentes con algunos cambios en los elementos de la situación experimental. Los resultados de la comparación entre la prueba por PCE y la prueba por cobertura de decisiones indican que no hay diferencia significativa en la efectividad de ambas pruebas. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y local dependiente del sitio. Sin embargo el resultado observado puede no ser real debido a que no se ha variado el protocolo experimental. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas para esta pieza y el poder de confirmación de los resultados es del 15%.

11.4.3 Generar Pieza de Conocimiento para Investigador

De manera similar a la actividad anterior, en esta última actividad se utiliza la tabla con información del análisis realizado (sección 11.3.3) y la tabla de variaciones (sección 11.4.1). A partir de estas tablas se empaqueta la pieza de conocimiento para el investigador. Para construir esta pieza de conocimiento se usa el formulario PCR-12 (apéndice D.12). En la

Información de la Partición			
Id Partición:	01.a		
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.		
Replicaciones de la partición:	5 (upm01..upm05)		
Resultados del análisis de similitudes:	Con Cambios		
Tipo de resultados encontrados:	Consistentes		
Información de los resultados:	Q=3.1, p=0.5348, I²=0%, T²= 0 P.C.E.: 69.69% [44.22; 95.15] f. observadas: 62.4, 81.45, 75.78, 66.6, 62.2 Cob. Dec.: 67.9% [35.45; 100] f. observadas: 57.59, 84.88, 70.27, 63.38, 63.39		
Tipo de inconsistencia observada:	inconsistencias no observadas		
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	5 (upm01..upm05)		
Comparación:	P.C.E. = Cob. Dec. (Sin diferencias en efectividad)		
Tipo de efecto global:	No significativo		
Tamaño del efecto Global:	0.0924 [-0.1136; 0.2985]		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	60%	Semi-objetivo	Nativa
Sitio:	0%	Local	Interna
Protocolo experimental:	0%	Artificial	-
Operacionalizaciones de los constructos:	13%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados	Operacional
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes	-
Poder de confirmación de resultados	15%		
Confiabilidad de la pieza de conocimiento:	Confiable		

Tabla 11.7: Pieza de conocimiento para investigador

tabla 11.7 se muestra la estructura principal del formulario PCR-12 con la información de la pieza de conocimiento resultante para el investigador.

De acuerdo a la información de esta pieza de conocimiento, el investigador obtiene los siguientes resultados. En esta partición seleccionada se observan resultados consistentes entre sí que permiten generar la pieza de conocimiento. De acuerdo a los resultados de la comparación de tratamientos, la prueba por PCE y la prueba por cobertura de decisiones indican que no hay diferencia significativa en la efectividad de ambas pruebas. El grado de verificación logrado en esta partición es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y global independiente del sitio. Los tipos de replicaciones identificados son: *nativa* al participar los investigadores del mismo grupo de investigación, *interna* al realizarse en el mismo sitio, y *operacional* al variarse las operacionalizaciones de los constructos. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 15%.

11.5 Resultados de la Evaluación

La tipología de replicaciones propuesta ha facilitado la identificación de varias replicaciones hechas de forma literal, nativa e interna elegidas para realizar comparaciones y obtener conocimiento de ellas. El *comparador de replicaciones* ha permitido conocer el grado de verificación alcanzado en este tipo de replicaciones tras comparar sus situaciones y resultados experimentales. Estas comparaciones han permitido generar conocimiento que se ha empaquetado en piezas que puedan servir tanto para el profesional como para el investigador.

Nótese que el estado de la práctica para llevar a cabo comparaciones entre replicaciones era que sólo se comparaban los resultados experimentales y se intentaba agregarlos a través de técnicas de meta-análisis. El *comparador de replicaciones* permite al investigador comparar las situaciones experimentales de varias replicaciones, comparar los resultados experimentales no sólo a través del meta-análisis sino también empleando otras técnicas que facilitan examinar las variaciones de los tratamientos. Los distintos tipos de comparaciones hechos en este sub-proceso permiten conocer el grado de verificación logrado en las replicaciones de una partición y empaquetar este conocimiento para que sea usado por el profesional o el investigador.

Capítulo 12

Evaluación con Replicaciones Diferenciadas-Nativas-Internas. Métrica Faltas Observadas

RESUMEN: En este capítulo se evalúan las ideas del marco para la comparación de replicaciones referentes a las comparaciones entre replicaciones. Esta evaluación se realiza a través del sub-proceso *comparador de replicaciones* usando una partición compuesta por una serie de replicaciones de tipo diferenciada, nativa e interna.

En este capítulo se evalúa el *comparador de replicaciones* con la partición 06.b. descrita en la sección 10.2.7. Esta partición se compone de diez replicaciones internas (*upm01..10*) realizadas en diez años consecutivos por el grupo de investigación GrISE de la FI-UPM. Aunque estas replicaciones se han realizado de forma interna, en varias replicaciones se han variado, principalmente, elementos de la situación experimental referentes a los experimentadores, protocolo experimental y poblaciones. En esta partición se compara la efectividad medida como el porcentaje de faltas observadas de las pruebas por cobertura de decisiones y lectura de código por abstracciones sucesivas.

Ya que en el capítulo anterior se ha generado la tabla maestra (en la sección 11.1) en este capítulo se omite esta actividad. A continuación se realizan los tres sub-procesos restantes del *comparador de replicaciones* utilizando la partición 06.b.

12.1 Sub-proceso Identificación de Similitudes

Como se describe en el procedimiento para comparar replicaciones, este sub-proceso se compone de dos actividades que son:

- Generar tablas de particiones
- Identificar similitudes

12.1.1 Generar Tablas de Particiones

Para generar las tablas de particiones se usa la información de la tabla maestra referente a las replicasiones *upm01..10* que conforman la partición 06.b. En la tabla 12.1 se muestra la tabla de partición referente a los experimentadores y sitio. Las tablas completas de esta partición se encuentran en el apéndice G.

12.1.2 Identificar Similitudes

En esta actividad del sub-proceso se utilizan las tablas de particiones antes descritas más el procedimiento de automatización (apéndice B) para generar las representaciones de similitudes de esta partición.

Siguiendo la automatización de cálculos, primeramente se organizan las tablas de particiones de las diferentes secciones de la situación experimental en un fichero de texto CSV¹ que representa la partición 06.b.

El fichero CSV de la partición 06.b se ingresa como parámetro al programa *Transpose* para generar ficheros de texto que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición. Los ficheros de texto resultantes se utilizan para generar de forma automática las representaciones de las similitudes. Este paso se realiza tras cargar el fichero `r_script01.txt` en el entorno R (R Dev. Core Team, 2011). Este fichero contiene varias funciones en R (R Dev. Core Team, 2011) para calcular las diferentes matrices de distancias así como las funciones para realizar los análisis de conglomerados jerárquico y de escalado multidimensional.

Una vez cargado en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`, se obtienen dos ficheros EPS que representan de forma gráfica las similitudes entre las replicasiones de la partición. En la figura 12.1 se muestran los gráficos resultantes que representan las similitudes de situaciones experimentales en esta partición².

12.2 Sub-proceso Análisis de Resultados Experimentales

En este sub-proceso se comparan los resultados experimentales de las replicasiones de la partición y junto con los cambios identificados en las situaciones experimentales se determina el grado de verificación logrado en la partición. El sub-proceso *análisis de resultados experimentales* se compone de las siguientes actividades:

- Generar tabla de resultados experimentales
- Realizar comparación de resultados experimentales
- Generar tabla con información del análisis realizado

12.2.1 Generar Tabla de Resultados Experimentales

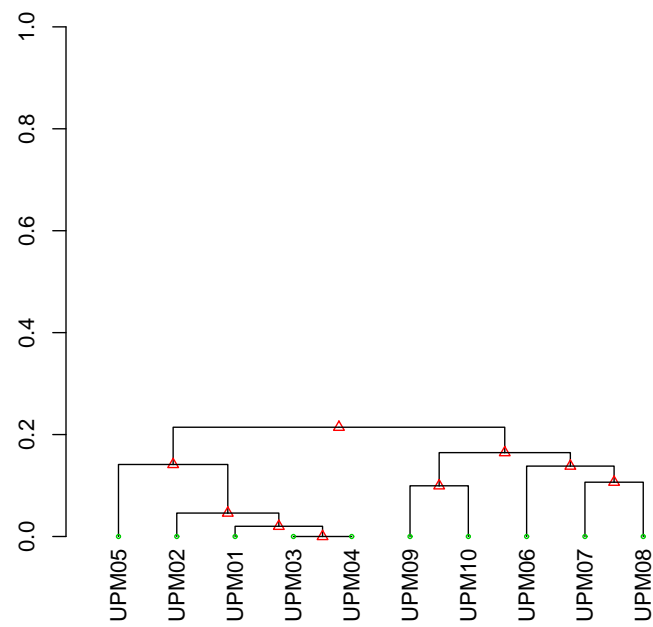
En esta actividad se extraen los resultados experimentales de las replicasiones que conforman la partición 06.b y se colocan en el formulario PCR-7 (apéndice D.7). En la tabla 12.2 se muestran los resultados experimentales obtenidos de la partición 06.b.

¹En el primer enlace de la sección *resources* de la url <http://www.osgg.net> se encuentra disponible el fichero de texto CSV de esta partición así como los ficheros generados de forma automática que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición.

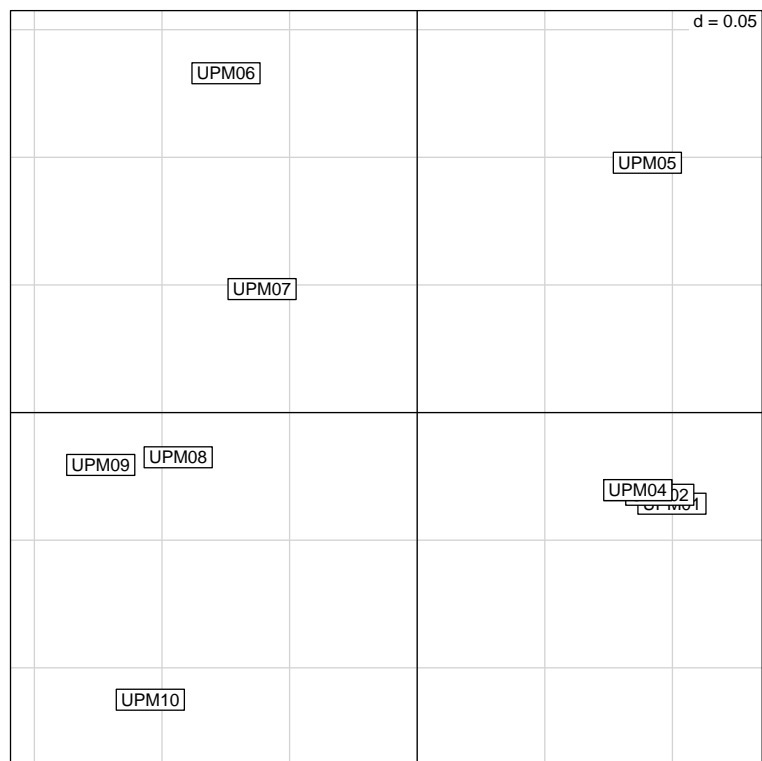
²Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`.

Estructura	Elemento principal	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10
Experimentadores	Diseñador	1.1	Vegas	X	X	X	X	X	X	X	X		
		1.1	Gómez									X	X
		1.1	Solari						X	X	X		
	Entrenador	1.2	Juristo	X	X	X	X	X	X	X	X	X	X
		1.2	Vegas	X	X	X	X		X	X	X	X	X
		1.2	Moreno					X	X				
	Monitor	1.3	Juristo	X	X	X	X			X	X	X	X
		1.3	Vegas	X	X	X	X	X	X	X	X	X	X
		1.3	Dieste		X								
		1.3	Ferre		X								
		1.3	Moreno					X	X				
		1.3	Gómez								X	X	X
	Medición	1.4	Vegas	X	X	X	X						X
		1.4	Gómez								X		
		1.4	Solari					X	X	X		X	
	Analista	1.5	Vegas	X	X	X	X	X					
		1.5	Gómez								X	X	X
		1.5	Solari						X	X			
Sitio		2	UPM	X	X	X	X	X	X	X	X	X	X

Tabla 12.1: Tabla de particiones referente a los experimentadores y sitio



(a) Dendrograma



(b) MDS

Fig. 12.1: Representación de similitudes, partición 06.b

Id exp.	Cob. Decisiones			Lec. Cód. por A.S.		
	Tamaño (n_1)	Porcentaje faltas observadas (\bar{y}_1)	Desv. est. (s_1)	Tamaño (n_2)	Porcentaje faltas observadas (\bar{y}_2)	Desv. est. (s_2)
upm01	42	57.59	19.56	42	47.27	18.79
upm02	39	84.88	19.56	39	34.06	21.39
upm03	29	70.27	19.19	29	46.3	18.95
upm04	36	63.38	18.64	36	55.55	23.6
upm05	32	63.39	23.5	32	54.91	25.16
upm06	46	35.51	20.37	46	53.62	23.02
upm07	48	40.97	23.56	48	52.78	21.01
upm08	40	28.75	19.24	48	36.81	18.5
upm09	50	38	24.06	49	57.48	19.26
upm10	25	31.33	21.14	25	58.66	23.13

Tabla 12.2: Resultados experimentales de la partición 06.b

12.2.2 Realizar Comparación de Resultados Experimentales

En esta actividad del sub-proceso se analizan los resultados experimentales de la partición. El análisis de los resultados se lleva a cabo a través del uso de gráficos de control y de técnicas estadísticas de agregación de resultados i.e. meta-análisis.

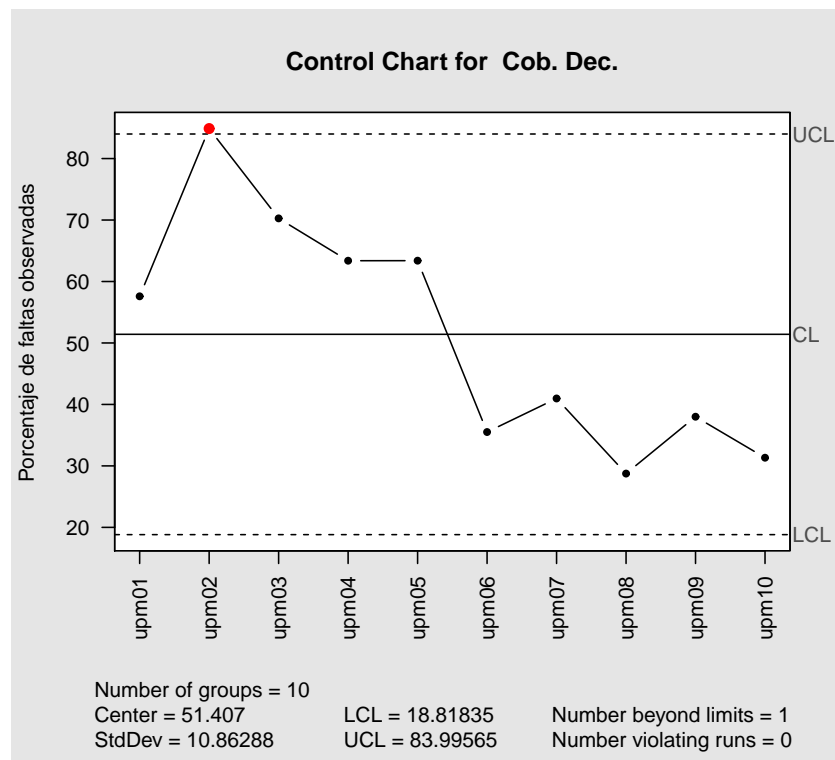
Para realizar estos análisis se usa la tabla de resultados experimentales de esta partición. A partir de esta tabla se generan dos gráficos de control, uno por cada tratamiento así como el meta-análisis que contiene entre otras cosas los tamaños de efecto estimados de las replications así como el tamaño de efecto global estimado de la partición. Estos análisis se llevan a cabo cargando el fichero `r_script02.txt` en el entorno R (R Dev. Core Team, 2011) por cada una de las particiones, en esta caso la partición 06.b. Este fichero recibe como entrada la tabla de resultados experimentales de la partición. Como salida se generan tres ficheros EPS, dos ficheros que contienen los gráficos de control del par de tratamientos (cob. de decisiones y lec. de código por abs. sucesivas), y otro fichero que contiene los resultados del meta-análisis. Con los gráficos de control y meta-análisis resultantes más las representaciones de similitudes de situaciones experimentales obtenidas en actividades previas se está en situación de decidir el grado de verificación que alcanzan los resultados experimentales.

En la figura 12.2 se muestran los gráficos de control de los dos tratamientos examinados en esta partición (cob. de decisiones y lec. de código por abs. sucesivas), mientras que en la figura 12.3 se muestran los resultados del meta-análisis³.

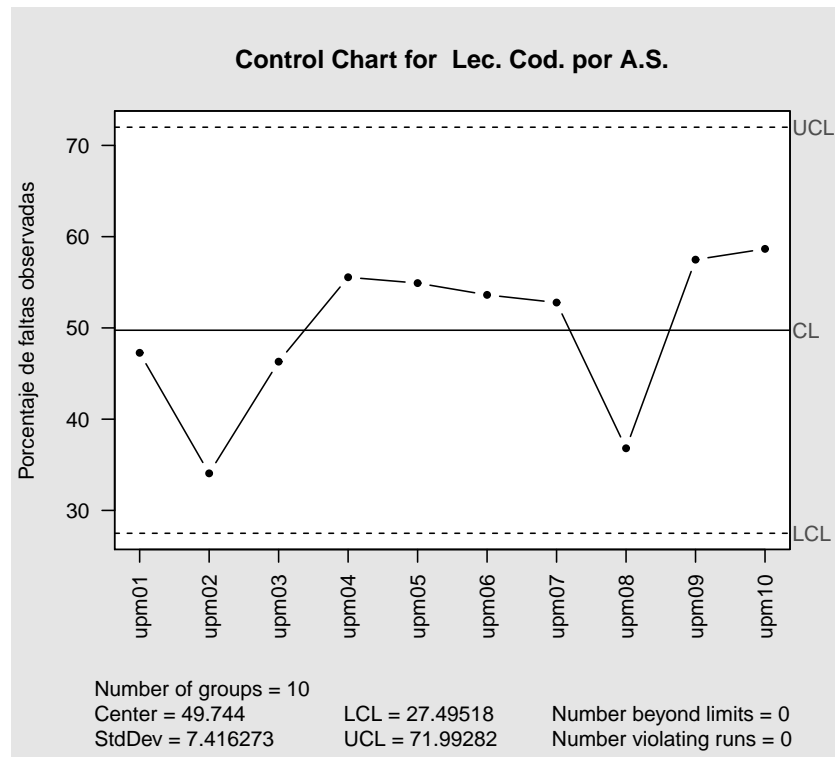
De acuerdo a los gráficos de control de las figuras 12.2a y 12.2b, en el gráfico que representa la prueba estructural por cobertura de decisiones (figura 12.2a) se observa que el valor de la replicación *upm02* se encuentra fuera de los límites de control, también en ese gráfico se observa una tendencia en los resultados siendo el porcentaje de faltas observadas en las replications *upm01* a *upm05* superior al 50% mientras que en las replications *upm06* a *upm10* este porcentaje es menor al 50%. Por otra parte en el gráfico de control que representa la prueba de lectura de código por abstracciones sucesivas (figura 12.2b) las observaciones se encuentran dentro de los límites de control.

Al examinar el meta-análisis de esta partición (figura 12.3) se observa que los tamaños de efecto estimados se distribuyen tanto en favor de la prueba por cobertura de decisiones (replications *upm01..05*) y viceversa (*upm06..upm10*). Esta contradicción en los tamaños de efecto de las replications de la partición ocasiona una heterogeneidad alta

³Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script02.txt`.



(a) Prueba por cob. de decisiones



(b) Prueba por lec. de cód. por abs. sucesivas

Fig. 12.2: Gráficos de control resultantes

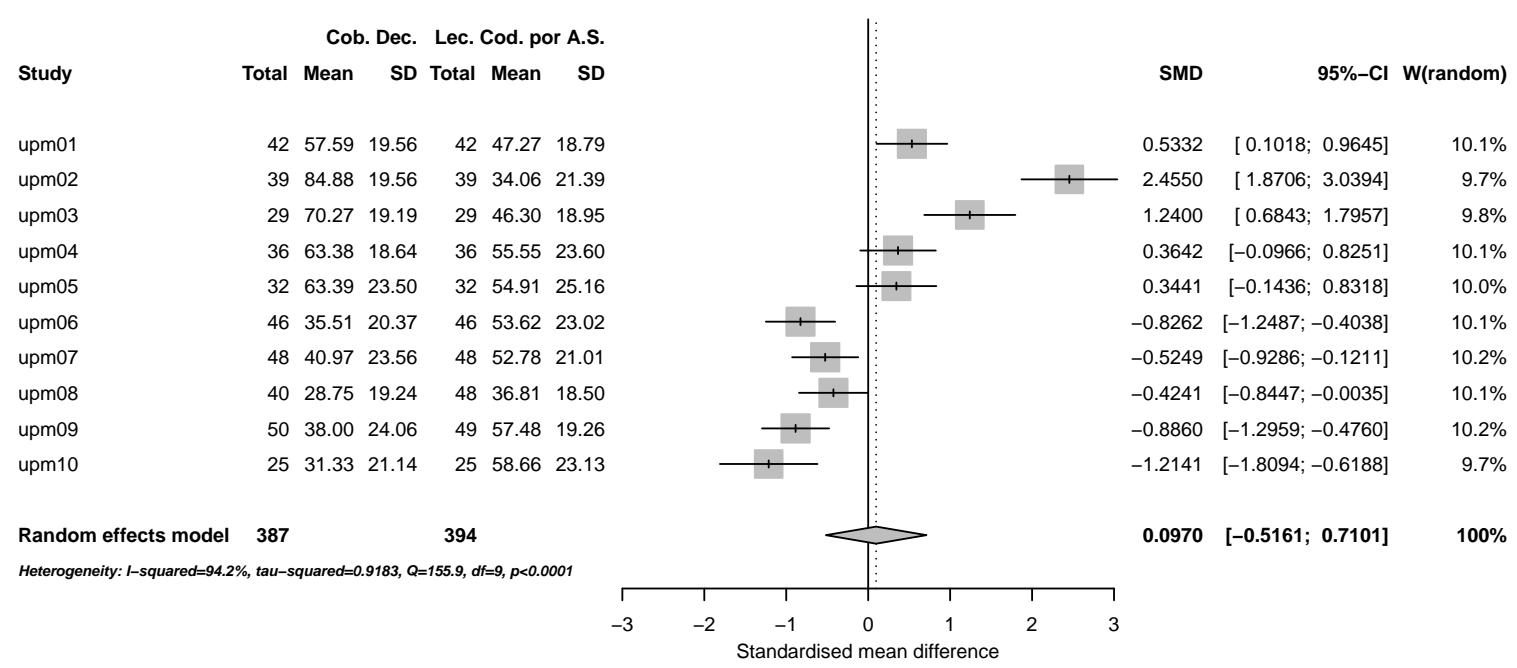


Fig. 12.3: Meta-análisis de las pruebas por cob. de decisiones y lec. de cód. por abs. sucesivas

en los resultados, como se observa en la prueba de heterogeneidad cuyo resultado es significativo ($Q = 155.9$, $p < 0.0001$), de manera similar los valores de los índices I^2 y T^2 indican una heterogeneidad alta. El índice I^2 indica un porcentaje de variabilidad total en el conjunto de tamaños de efecto estimados del 94.2%, por otra parte el índice T^2 que representa la varianza inter-estudios contiene un valor elevado cercano a uno⁴.

Respecto a las figuras 12.1a y 12.1b que muestran los resultados de la identificación de similitudes se observa en el dendrograma y en el MDS dos agrupaciones de replicaciones, en un grupo se encuentran las replicaciones *upm01* a *upm05* mientras que en el otro las replicaciones *upm06* a *upm10*. En el caso del dendrograma la diferencia entre ambos grupos se produce a una distancia poco superior de 0.2 esto indica que la mayoría de los elementos de la situación experimental son comunes en las replicaciones de la partición.

De acuerdo a los resultados presentados en los gráficos de control, meta-análisis e identificación de similitudes, se tienen resultados inconsistentes producto de uno de los gráficos de control y del meta-análisis mientras que en los resultados de la identificación de similitudes se observan cambios en la situación experimental. Si se emplean todos las replicaciones de esta partición para obtener conocimiento, éste no sería confiable debido a las inconsistencias observadas. En esta situación, el investigador decide si intenta o no obtener conocimiento a partir de agrupar resultados consistentes en la partición.

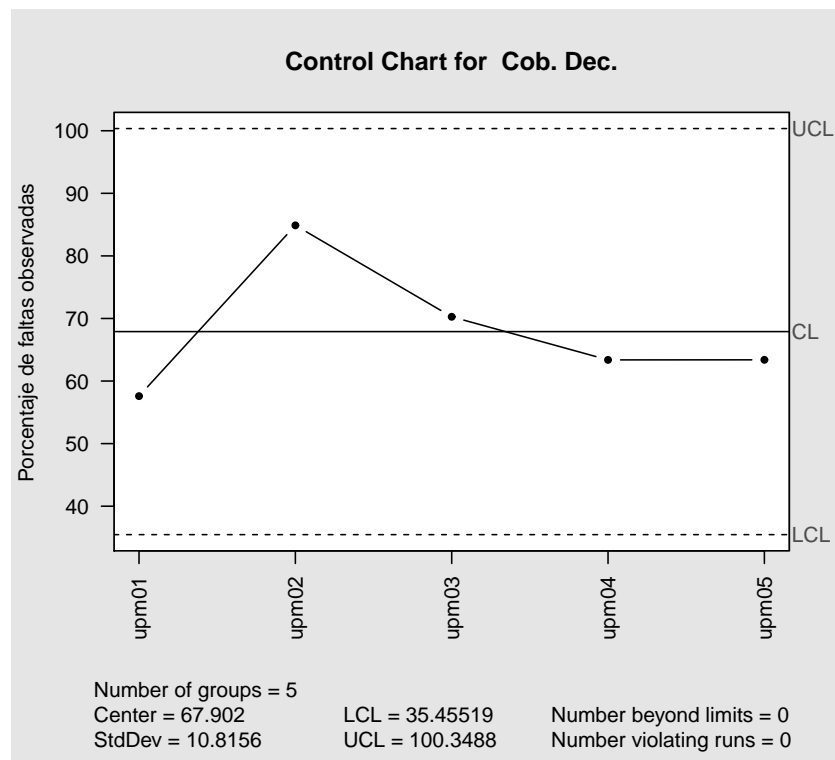
Aunque los resultados obtenidos en esta partición son inconsistentes, se ha decidido continuar con esta actividad del sub-proceso e intentar obtener algún conocimiento. Siguiendo con esta actividad, a continuación se examina la distribución de los tamaños de efecto tanto en los gráficos de control como en el meta-análisis. En la distribución de resultados del gráfico de control de la prueba por cobertura de decisiones así como en la distribución de tamaños de efecto del meta-análisis se observan dos grupos de resultados. El primer grupo se compone de las replicaciones *upm01* a *upm05* mientras que el segundo grupo se compone de las replicaciones *upm06* a *upm10*. Con cada uno de estos grupos se vuelven a generar los gráficos de control y el meta-análisis para intentar obtener resultados consistentes que permitan empaquetar posteriormente una o varias piezas de conocimiento. A continuación se examina cada uno de los grupos de resultados identificados.

12.2.2.1 Análisis del Primer Grupo de Resultados

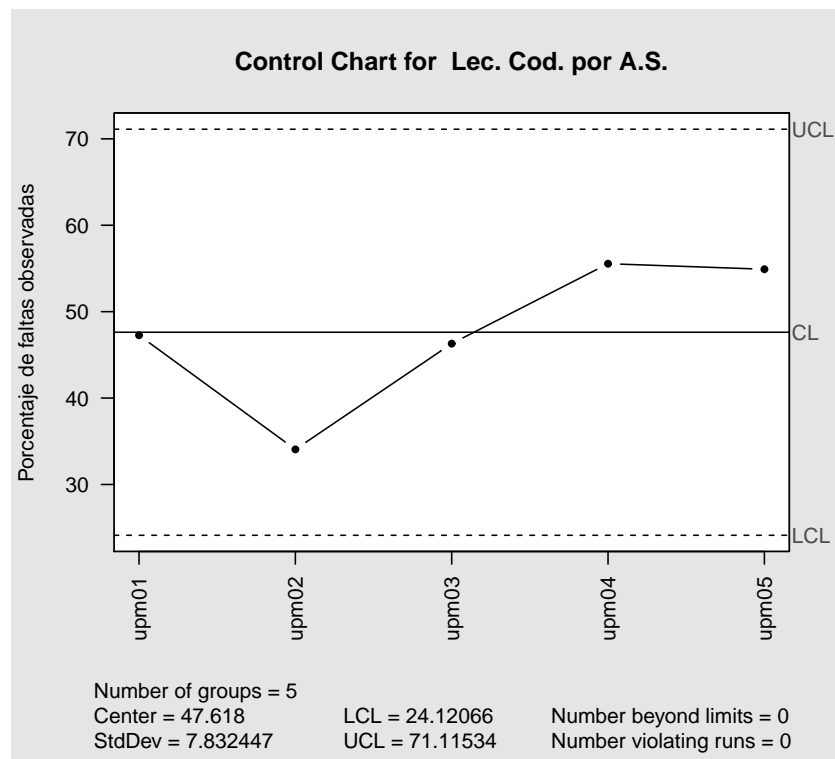
En las figuras 12.4 y 12.5 se muestran los gráficos de control y meta-análisis resultante del primer grupo de resultados,

En este primer grupo de resultados, los gráficos de control en ambos tratamientos (prueba por cob. decisiones y prueba por lec. de código por abstracciones sucesivas) se encuentran dentro de los límites de control, sin embargo en el meta-análisis se observa una heterogeneidad en los resultados debido en apariencia a la replicación *upm02*. En este grupo de replicaciones la prueba de heterogeneidad es significativa con un valor $Q = 41.3$ y $p < 0.0001$, el valor del índice I^2 indica un porcentaje de variabilidad del 90.3% mientras que el índice T^2 un valor de 0.6014. En este grupo se aparta el resultado de la replicación *upm02* y se usa el resto de resultados de este grupo para obtener resultados consistentes y conocer el grado de verificación logrado en este grupo. Con estas cuatro replicaciones (*upm01*, *upm03*, *upm4* y *upm05*) se vuelven a generar los gráficos de control y meta-análisis. En las figuras 12.6 y 12.7 se muestran los resultados del meta-análisis y gráficos

⁴A diferencia de los resultados consistentes observados en el capítulo anterior 11 donde el valor obtenido en T^2 es igual a cero, i.e. la varianza inter-estudios es igual a cero (ver sección B.8).



(a) Prueba por cob. de decisiones



(b) Prueba por lec. de cód. por abs. sucesivas

Fig. 12.4: Gráficos de control del primer grupo de resultados

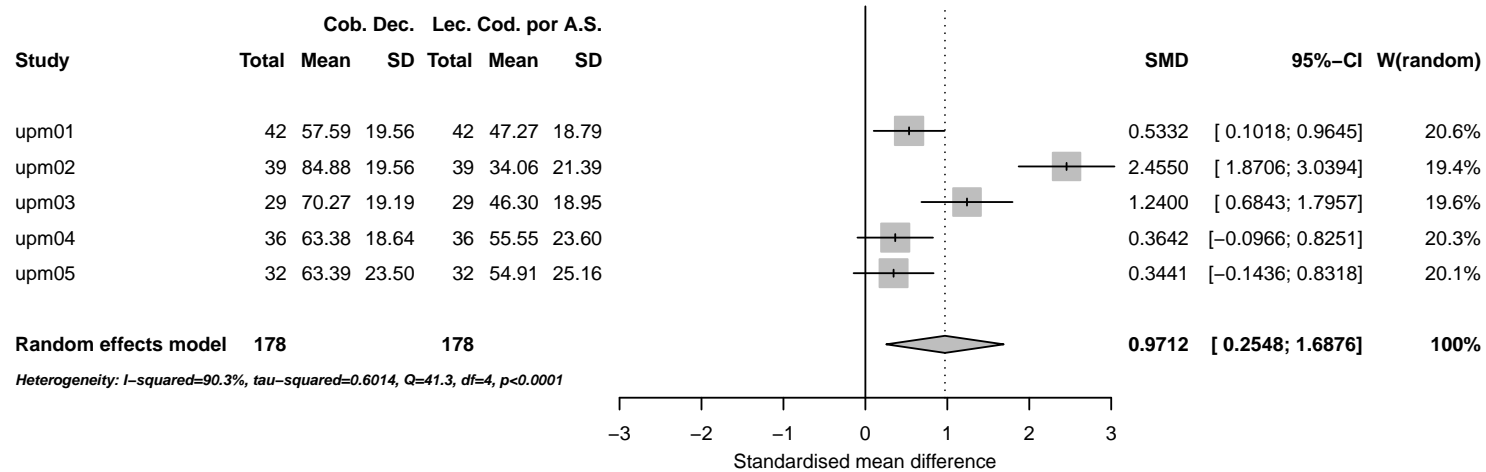


Fig. 12.5: Meta-análisis de las pruebas cob. de decisiones y lec. de cód. por abs. sucesivas del primer grupo de resultados

de control tras apartar *upm02*.

Como se observa en los gráficos de control de las figuras 12.6a y 12.6b, los resultados de las cuatro replicaciones se encuentran dentro de los límites de control, sin embargo en el meta-análisis se observa una heterogeneidad media ($I^2 = 58.3\%$) que en apariencia se debe a *upm03*. Esta replicación se aparta para obtener resultados consistentes con el resto de replicaciones (*upm01*, *upm04* y *upm05*). En las figuras 12.8 y 12.9 se muestran los gráficos de control y meta-análisis resultante tras apartar *upm03*.

Como se observa en las figuras 12.8a, 12.8b los resultados en ambos gráficos de control se encuentran dentro de los límites de control, en el caso de la figura 12.9 se observa una homogeneidad en los tamaños de efecto estimados por lo que no se observa alguna heterogeneidad en los resultados. Ambos resultados tanto los gráficos de control como meta-análisis son consistentes y pueden usarse para conocer el grado de verificación lo grado.

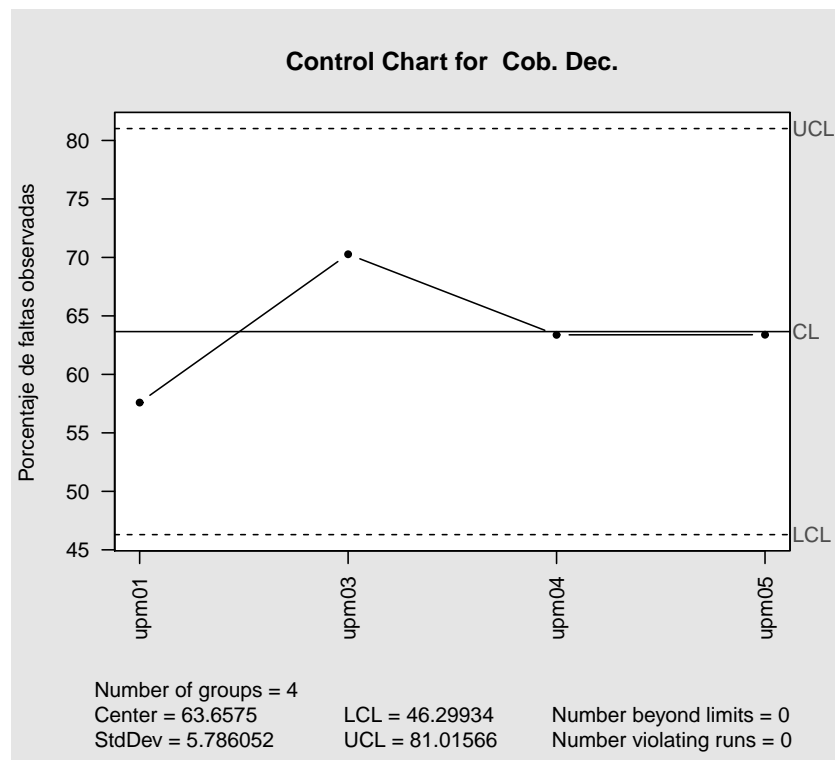
12.2.2.2 Análisis del Segundo Grupo de Resultados

Una vez que se identificaron resultados consistentes en el anterior grupo, a continuación se examina el segundo grupo de resultados compuesto por las replicaciones *upm06* a *upm10*. En las figuras 12.10 y 12.11 se muestran los gráficos de control y meta-análisis de este segundo grupo de resultados.

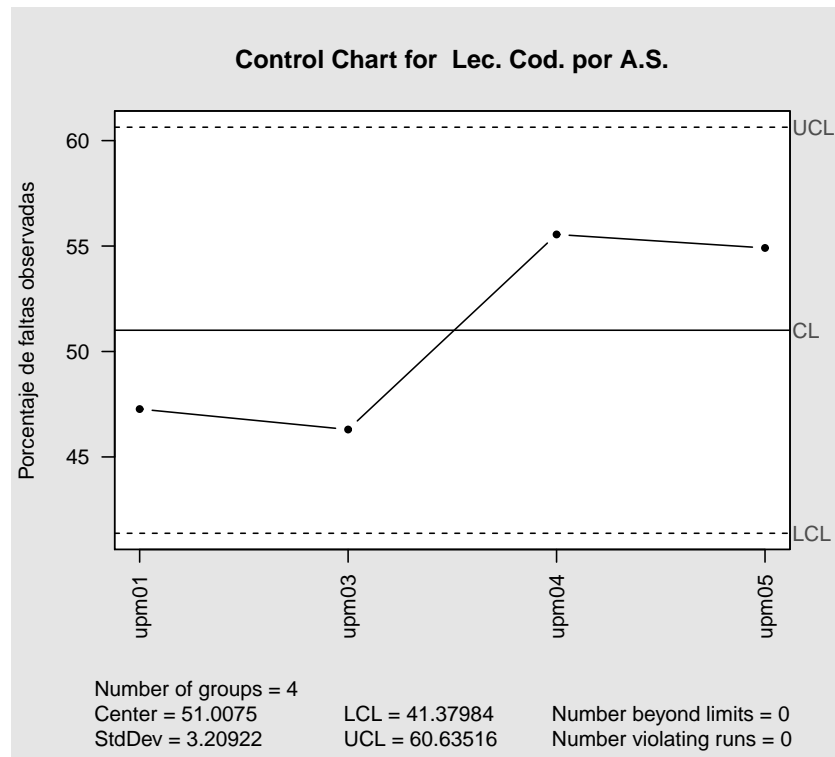
Como se muestra en las figuras 12.10a y 12.10b los resultados de las replicaciones se encuentran para cada tratamiento dentro de los límites de control, no obstante, en el meta-análisis (figura 12.11) se observa cierta heterogeneidad en los resultados, que aunque la prueba de heterogeneidad no es significativa, el índice I^2 indica un porcentaje de variabilidad del 36.7% ocasionada en apariencia por *upm10*. Esta replicación es separada para obtener resultados consistentes en el resto de replicaciones (*upm06* a *upm10*). En las figuras 12.12 y 12.13 se muestran los gráficos de control y meta-análisis tras apartar *upm10*.

Tras realizar de nuevo los gráficos de control y meta-análisis, los resultados en ambos gráficos (figuras 12.12a y 12.12b) se encuentran dentro de los límites de control, mientras que en el meta-análisis se ha reducido el nivel de heterogeneidad a un nivel aceptable para considerar homogeneidad de resultados ($I^2 = 11.7\%$). Al obtener en estas replicaciones (*upm06* a *upm09*) resultados consistentes es posible conocer el grado de verificación alcanzado en este segundo grupo.

Cuando se intenta obtener conocimiento a partir de resultados inconsistentes como sucede en esta partición, es recomendable realizar un análisis de causas para tratar de identificar uno o varios elementos de la situación experimental que se cree influyen en los resultados. Continuando con esta actividad del sub-proceso se utilizan diagramas de causa y efecto (Ishikawa, 1990) para realizar el análisis de causas. Para obtener los dos conjuntos de resultados consistentes de esta partición primero se identificaron dos grupos de resultados inconsistentes (primer grupo *upm01..05* y segundo grupo *upm06..10*), posteriormente en cada grupo se identificaron otros resultados inconsistentes (*upm02* y *upm03* en el primer grupo, y *upm10* en el segundo grupo). Para el análisis de causas primero se examinan las inconsistencias intra-grupos y posteriormente se examinan las inconsistencias inter-grupos.



(a) Prueba por cob. de decisiones



(b) Prueba por lec. de cód. por abs. sucesivas

Fig. 12.6: Gráficos de control del primer grupo de resultados sin upm02

Study	Cob. Dec.			Lec. Cod. por A.S.		
	Total	Mean	SD	Total	Mean	SD
upm01	42	57.59	19.56	42	47.27	18.79
upm03	29	70.27	19.19	29	46.30	18.95
upm04	36	63.38	18.64	36	55.55	23.60
upm05	32	63.39	23.50	32	54.91	25.16
Random effects model	139			139		

Heterogeneity: $I^2=58.3\%$, $\tau^2=0.0841$, $Q=7.2$, $df=3$, $p=0.0657$

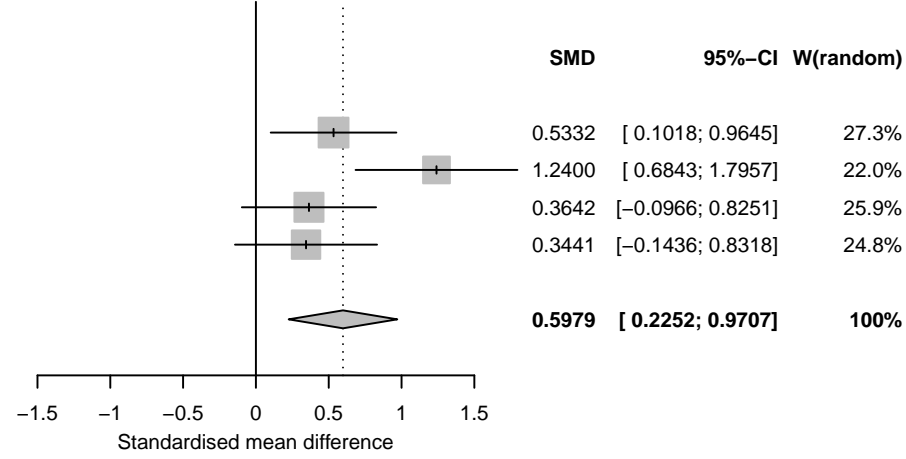
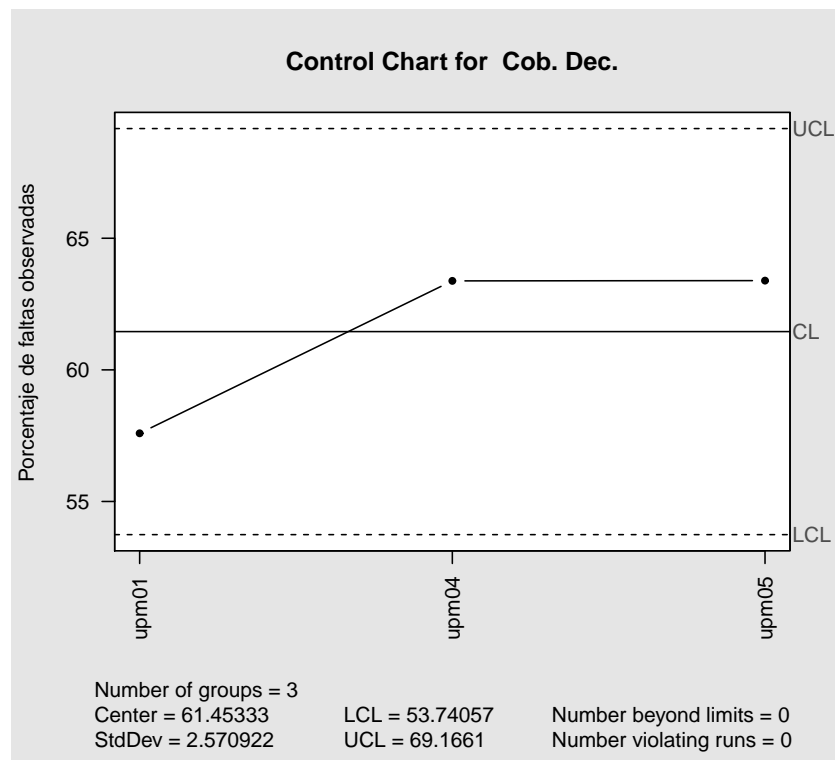
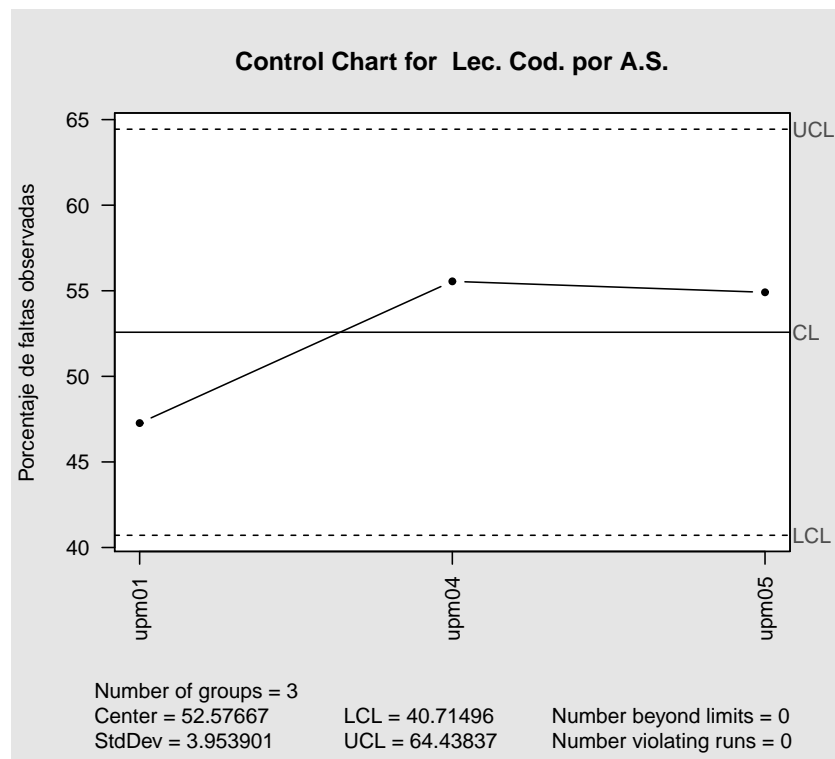


Fig. 12.7: Meta-análisis de las pruebas cob. de decisiones y lec. de cód. por abs. sucesivas del primer grupo de resultados sin upm02



(a) Prueba por cob. de decisiones



(b) Prueba por lec. de cód. por abs. sucesivas

Fig. 12.8: Gráficos de control del primer grupo de resultados sin upm02 y upm03

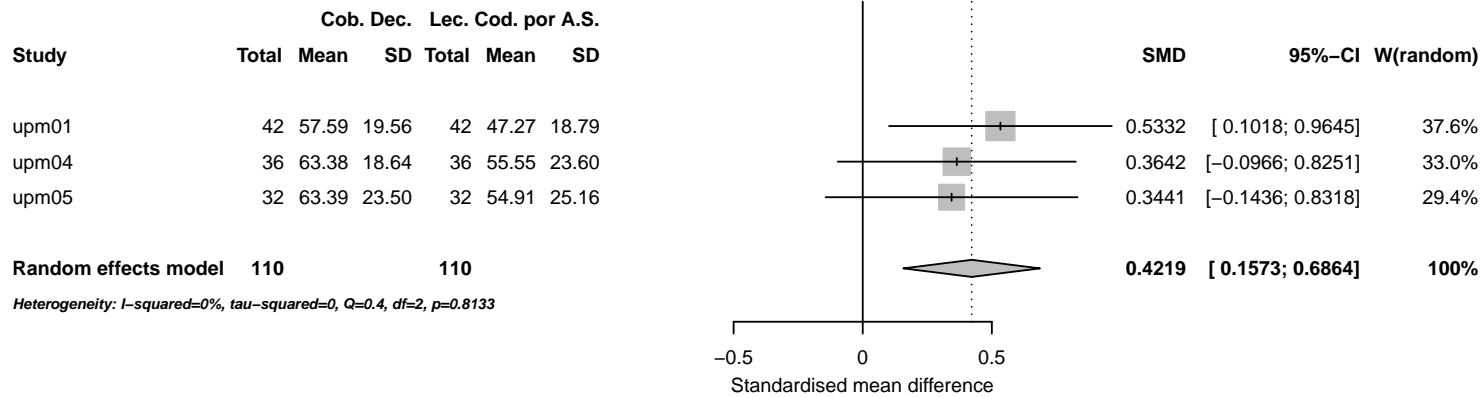
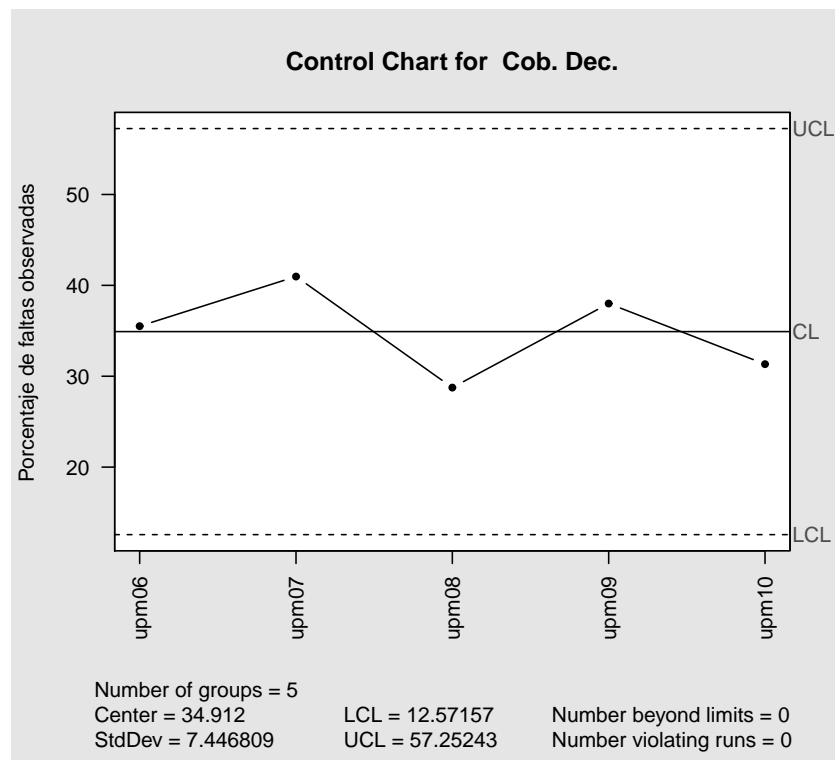
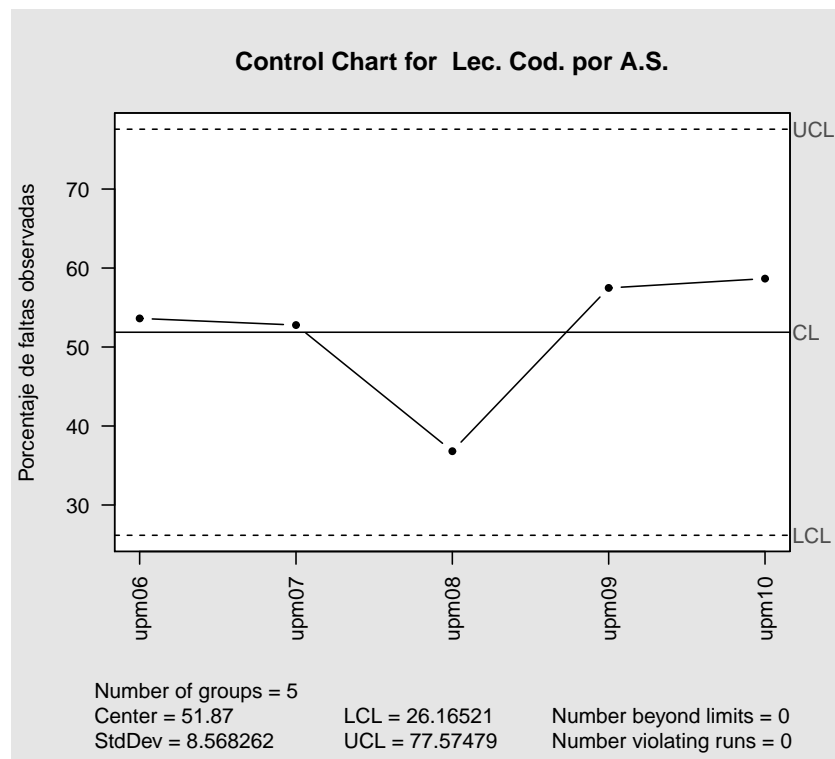


Fig. 12.9: Meta-análisis de las pruebas cob. de decisiones y lec. de cód. por abs. sucesivas del primer grupo de resultados sin upm02 y upm03



(a) Prueba por cob. de decisiones



(b) Prueba por lec. de cód. por abs. sucesivas

Fig. 12.10: Gráficos de control del segundo grupo de resultados

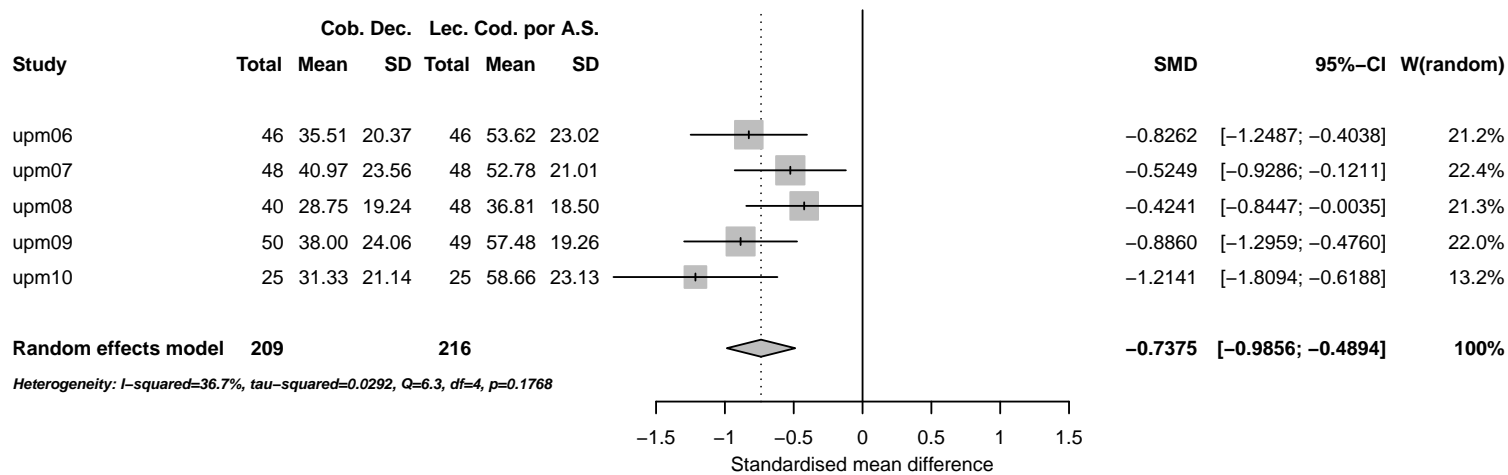
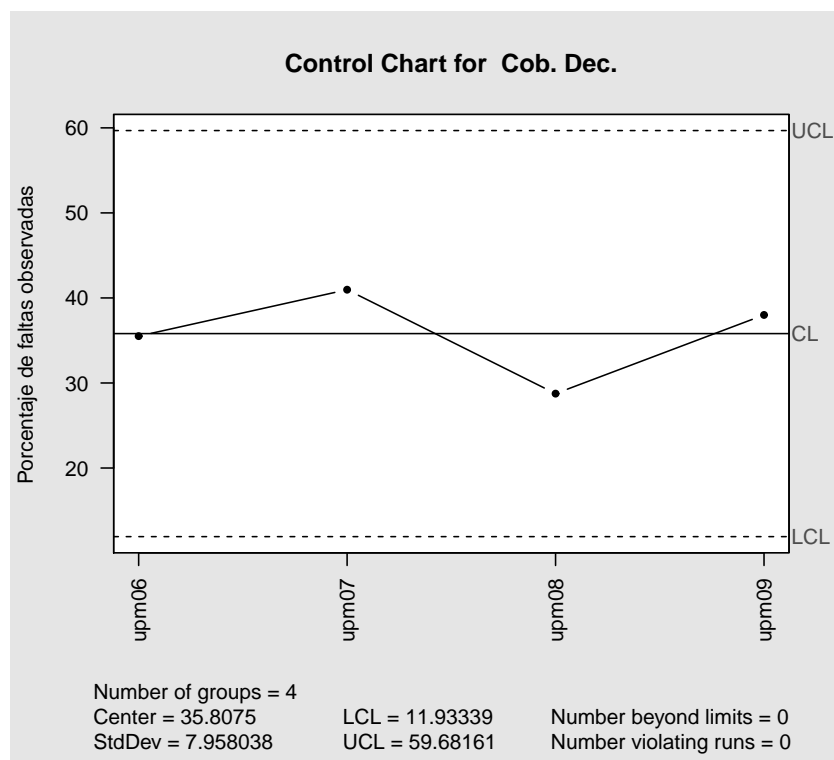
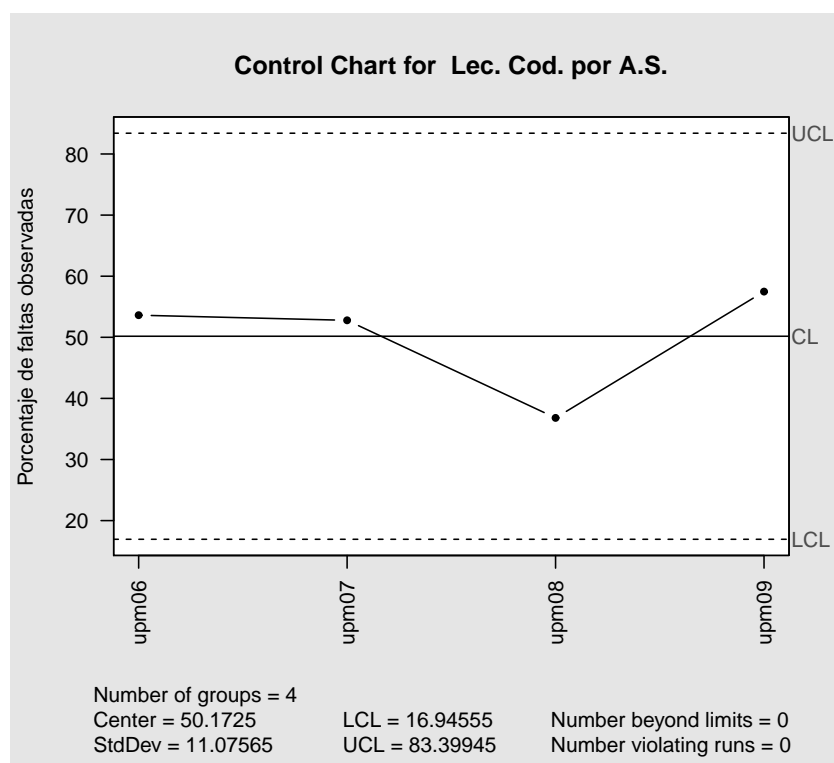


Fig. 12.11: Meta-análisis de las pruebas cob. de decisiones y lec. de cód. por abs. sucesivas del segundo grupo de resultados



(a) Prueba por cob. de decisiones



(b) Prueba por lec. de cód. por abs. sucesivas

Fig. 12.12: Gráficos de control del segundo grupo de resultados sin upm10

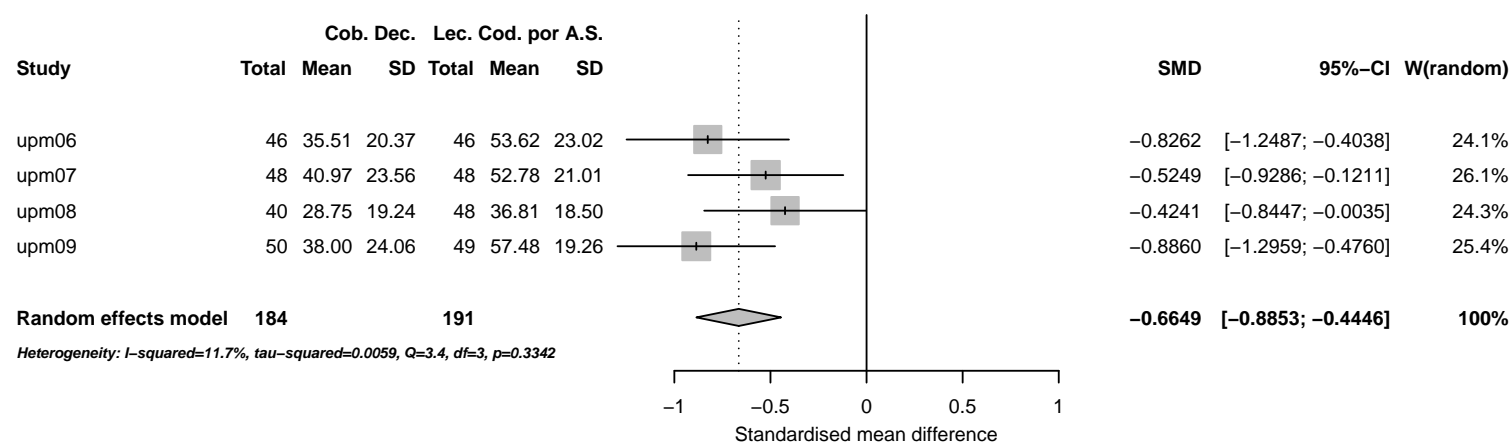


Fig. 12.13: Meta-análisis de las pruebas cob. de decisiones y lec. de cód. por abs. sucesivas del segundo grupo de resultados sin upm10

12.2.2.3 Análisis de Causas del Primer Grupo de Resultados

En las figuras 12.14 y 12.15 se muestran dos diagramas de Ishikawa (Ishikawa, 1990) con el número de cambios identificados en los elementos de la situación experimental del primer grupo de resultados (replicaciones *upm01* a *upm05*). En la tabla 12.3 se muestra el fragmento de las tablas de particiones con los cambios identificados.

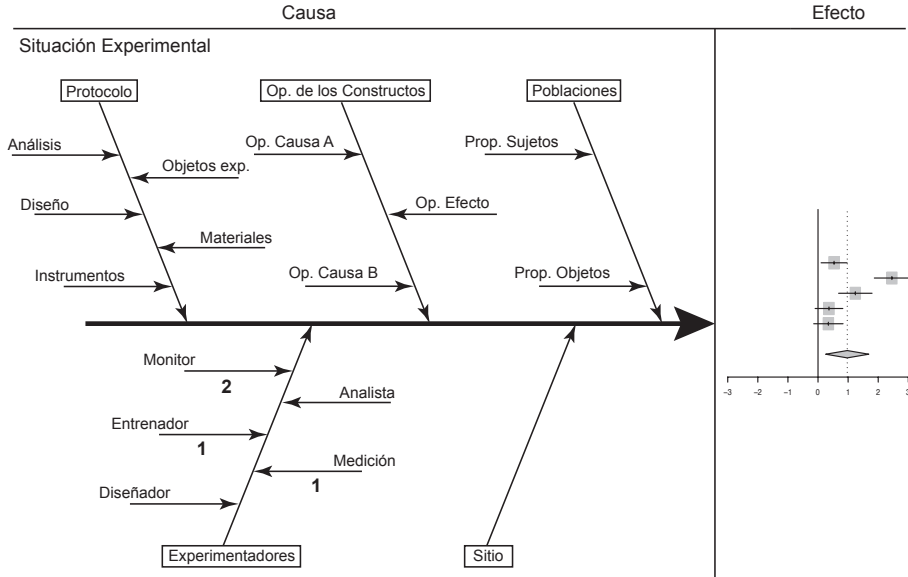


Fig. 12.14: Cambios identificados en experimentadores y sitio

Elemento		ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05
Entrenador		1.2	Juristo	X	X	X	X	X
		1.2	Vegas	X	X	X	X	
		1.2	Moreno					X
Monitor		1.3	Juristo	X	X	X	X	
		1.3	Vegas	X	X	X	X	X
		1.3	Dieste		X			
		1.3	Ferre		X			
		1.3	Moreno					X
Medición		1.4	Vegas	X	X	X	X	
		1.4	Solari					X
Cob. Dec.	Aspectos de transmisión del tratamiento	4.1.1.5	Con experiencia en curso universitario	X	X	X	X	X
		4.1.1.5	Sin experiencia en curso universitario	X				
Lec. Cód. A.S.	Aspectos de transmisión del tratamiento	4.1.2.5	Con experiencia en curso universitario	X	X	X	X	X
		4.1.2.5	Sin experiencia en curso universitario	X				

Tabla 12.3: Cambios identificados en los elementos de la situación experimental

A continuación se analizan los cambios identificados en los elementos de la situación experimental, en este caso el análisis comienza con el elemento *aspectos de transmisión del tratamiento* de las operacionalizaciones de los constructos que es uno de los elementos que tiene un sólo cambio. Respecto a este elemento, a diferencia del resto de replicaciones en *upm01* participaron dos investigadores durante la sesión del entrenamiento uno de ellos con experiencia para transmitir el tratamiento mientras que el otro en ese entonces carecía

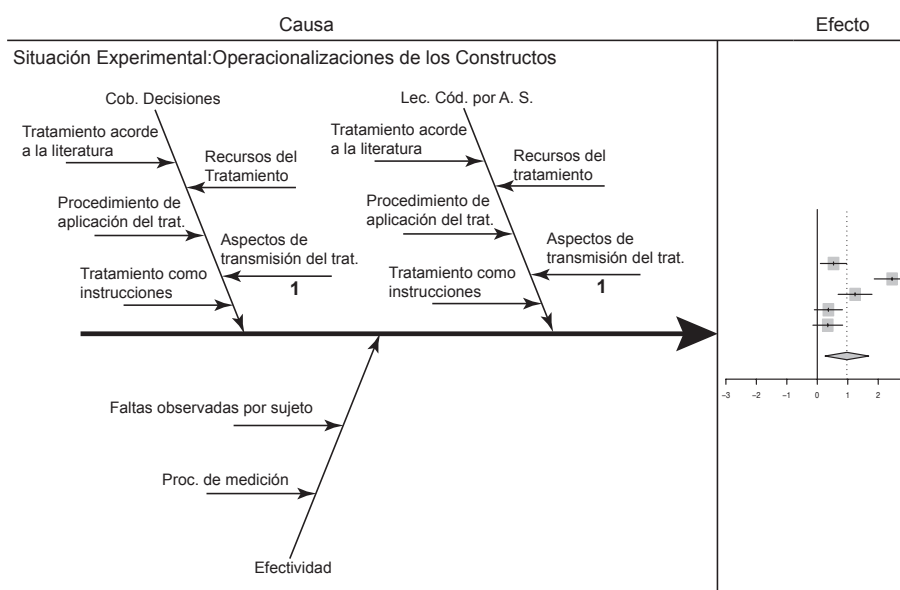


Fig. 12.15: Cambios identificados en operacionalizaciones de los constructos

de experiencia, en el resto de replicaciones, los investigadores cuentan con experiencia para transmitir el tratamiento a los sujetos. El cambio identificado en *upm01* parece no influir en los resultados, por ejemplo, se verifican los resultados entre las replicaciones *upm01* y *upm04* (figura 12.5). En este par de replicaciones sólo se varía este elemento de la situación experimental.

Analizando los elementos referentes a experimentadores, en la replicación *upm05* se observan cambios en los roles de entrenador, monitor y medición respecto a las replicaciones *upm01*, *upm03* y *upm04* que utilizan los mismos roles. Los cambios en los roles en *upm05* parecen no influir ya que se verifican los resultados con las replicaciones *upm01* y *upm04*.

En el caso de *upm02* los cambios en el rol de monitor parecen influir en los resultados. Esta replicación es más similar a *upm03*. A diferencia de *upm03*, en *upm02* participaron otros dos investigadores en el rol de monitores. Tal vez estos investigadores dieron algunas instrucciones a los sujetos que influyeron en los resultados de la replicación.

Por otra parte, *upm03* es igual a la replicación *upm04* sin embargo los resultados de *upm03* difieren del resto. En esta replicación otra condición no identificada pudo influir en los resultados.

De los cambios anteriormente discutidos se cree que el rol de monitor influye en los resultados. Este elemento puede ser testado en futuras replicaciones para verificar esta hipótesis.

12.2.2.4 Análisis de Causas del Segundo Grupo de Resultados

En las figuras 12.16, 12.17, 12.18 y 12.19 se usan los diagramas de Ishikawa (Ishikawa, 1990) para representar los cambios identificados en las replicaciones (*upm06* a *upm10*) de este segundo grupo de resultados. En la tabla 12.4 se muestra un extracto de las tablas de particiones con los cambios identificados.

A diferencia de los cambios identificados en el primer grupo de resultados, en ese

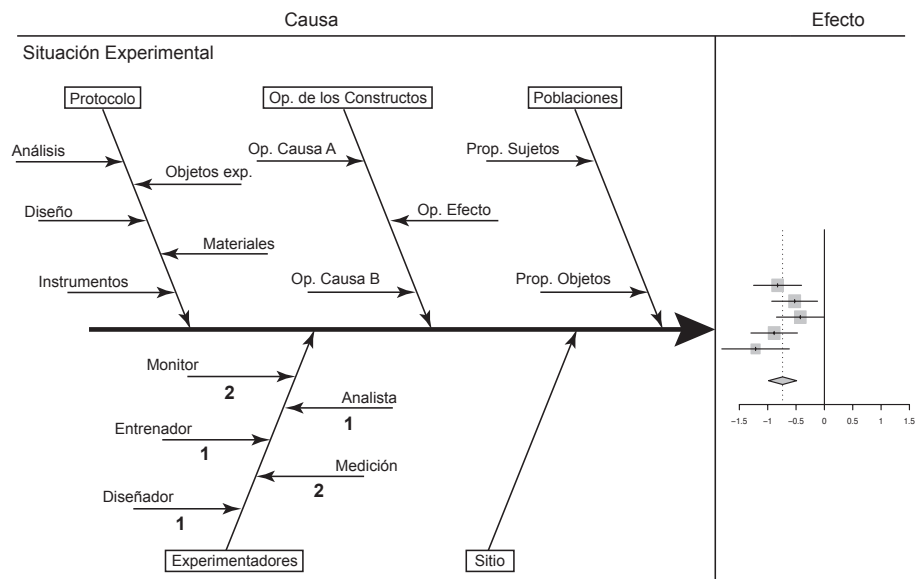


Fig. 12.16: Cambios identificados en experimentadores y sitio

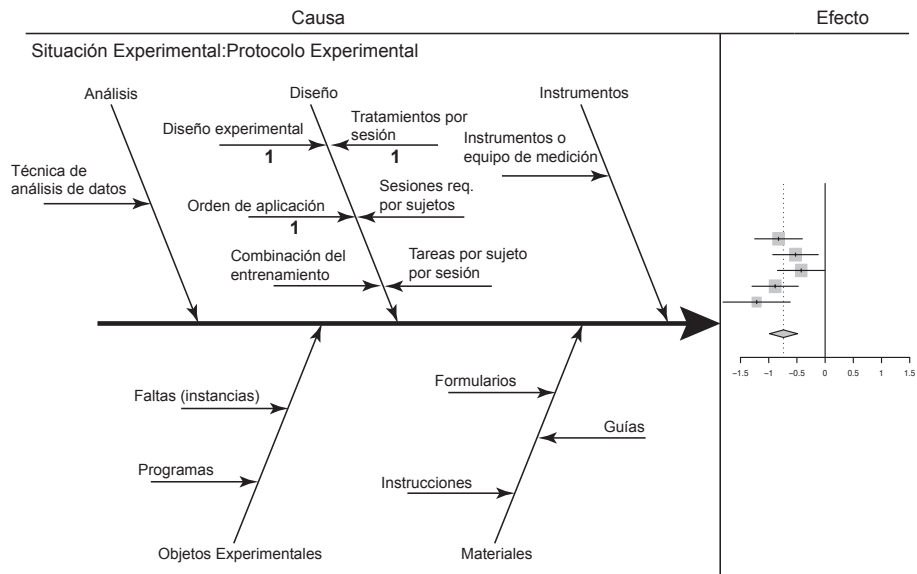


Fig. 12.17: Cambios identificados en el protocolo experimental

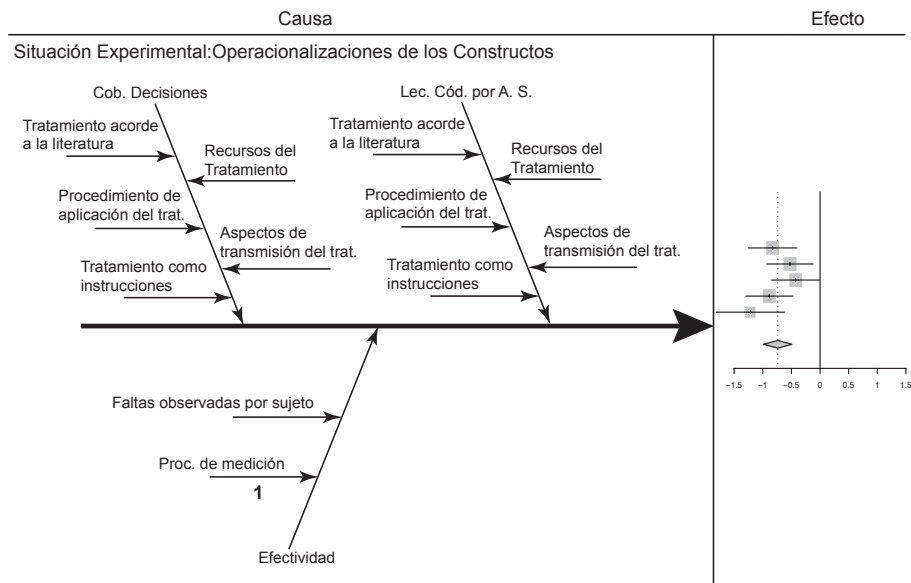


Fig. 12.18: Cambios identificados en operacionalizaciones de los constructos

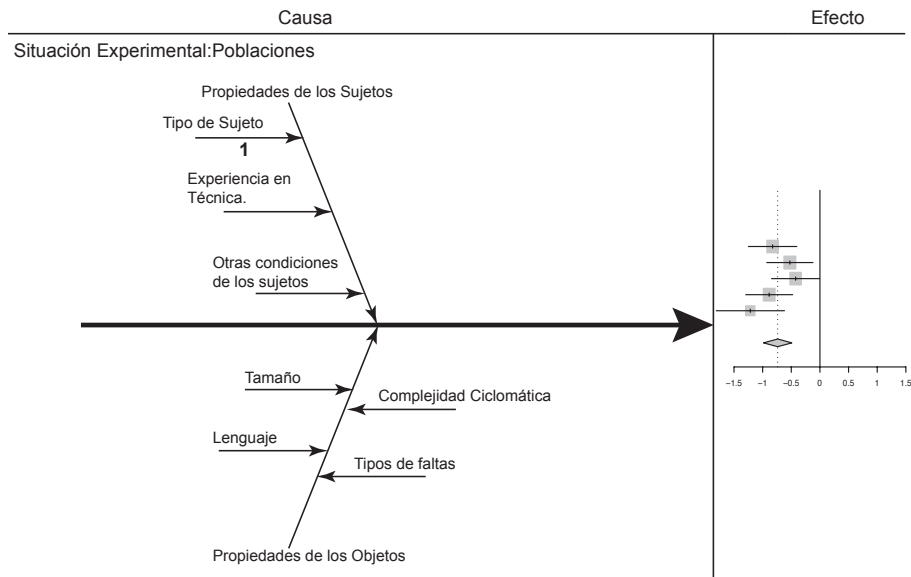


Fig. 12.19: Cambios identificados en poblaciones

Elemento	ID	Categoría	UPM06	UPM07	UPM08	UPM09	UPM10
Diseñador	1.1	Vegas	X	X	X		
	1.1	Gómez				X	X
	1.1	Solari	X	X	X		
Entrenador	1.2	Juristo	X	X	X	X	X
	1.2	Vegas		X	X	X	X
	1.2	Moreno	X				
Monitor	1.3	Juristo		X	X	X	X
	1.3	Vegas	X	X	X	X	X
	1.3	Moreno	X				
	1.3	Gómez			X	X	X
Medición	1.4	Vegas					X
	1.4	Gómez			X		
	1.4	Solari	X	X		X	
Analista	1.5	Gómez			X	X	X
	1.5	Solari	X	X			
Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X		
	3.2.1	Diseño Factorial				X	X
Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X		
	3.2.2	No Aplica (c/ suj app 1 tec)				X	X
	3.2.4	3 Técnicas, 1 Programa	X	X	X		
Tratamientos por sesión	3.2.4	3 Técnicas, 3 Programas				X	X
Procedimiento de medición (criterio de corrección)	4.2.2	Criterio relajado	X	X		X	X
	4.2.2	Criterio estricto			X		
Tipo de sujeto	5.1.1	Pregrado avanzado	X	X	X	X	X
	5.1.1	Grado (Máster o Doc.)					X

Tabla 12.4: Cambios identificados en los elementos de la situación experimental

segundo grupo se observan más cambios en los elementos de la situación experimental. Esto dificulta tratar de identificar posibles elementos que se cree influyen en los resultados.

En este segundo grupo, los resultados de *upm10* ocasionan un aumento en la heterogeneidad de esta partición (figura 12.11). A diferencia del resto de replicaciones, en *upm10* participaron estudiantes de pregrado avanzado y estudiantes graduados cursando estudios de Máster o Doctorado. Se cree que este cambio pudo ocasionar un aumento promedio en la efectividad de la prueba de lectura de código por abstracciones sucesivas. Otra variación en los tamaños de efecto estimados se observa entre los pares de replicaciones *upm06*, *upm09* y *upm07*, *upm08*. Sin embargo, entre estos dos pares de replicaciones no se observa algún cambio aparente que permita establecer alguna hipótesis inicial. Puede ser que un conjunto de los cambios identificados sean las causas de estas diferencias.

De los cambios identificados en este segundo grupo de resultados se cree que el tipo de sujeto pudo influir en los resultados. Esta hipótesis inicial puede ser testada en replicaciones posteriores.

12.2.2.5 Análisis de Causas Inter-grupos de Resultados

Una vez analizados los cambios a nivel intra-grupos, a continuación se realiza el análisis de causas inter-grupos. En las figuras 12.20, 12.21 y 12.22 se muestran los diagramas de Ishikawa (Ishikawa, 1990) con el número de cambios identificados entre estos dos grupos de resultados, mientras que en la tabla 12.5 se muestran los elementos variados.

Los cambios identificados en los elementos de la situación experimental entre estos dos grupos de resultados se relacionan con las instancias de faltas de los objetos experimentales, con los tipos de faltas descritos en las poblaciones y con uno de los elementos de las operacionalizaciones de causa de la prueba por cobertura de decisiones. En las replicaciones *upm01* a *upm05* se usaron dos versiones de cada programa. Cada versión

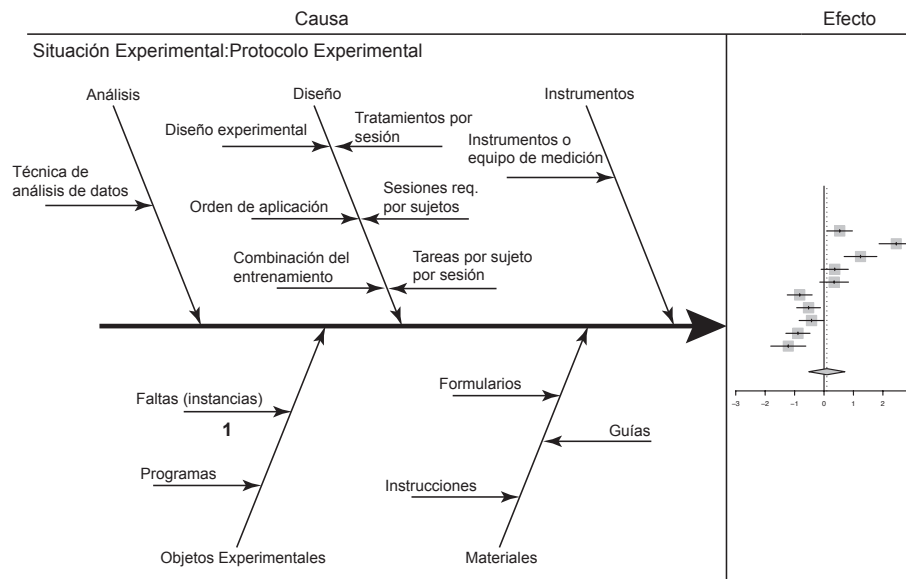


Fig. 12.20: Cambios identificados en protocolo experimental

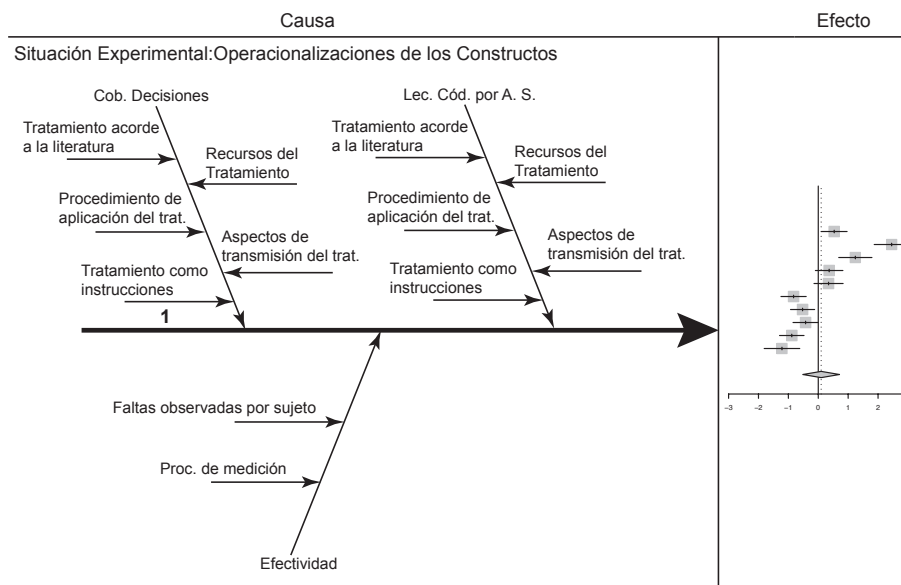


Fig. 12.21: Cambios identificados en operacionalizaciones de los constructos

Elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10
Faltas (Instancias)	3.4.2	Instancia G	X	X	X	X	X					
	3.4.2	Instancia H	X	X	X	X	X					
	3.4.2	Instancia I						X	X	X	X	X
Tratamiento como instrucciones	4.1.1.3	Cobertura de decisiones v2	X	X	X	X	X					
	4.1.1.3	Cobertura de decisiones v3						X	X	X	X	X
Tipos de Faltas	5.2.4	Grupo de Faltas G	X	X	X	X	X					
	5.2.4	Grupo de Faltas H	X	X	X	X	X					
	5.2.4	Grupo de Faltas I						X	X	X	X	X

Tabla 12.5: Cambios identificados en los elementos de la situación experimental

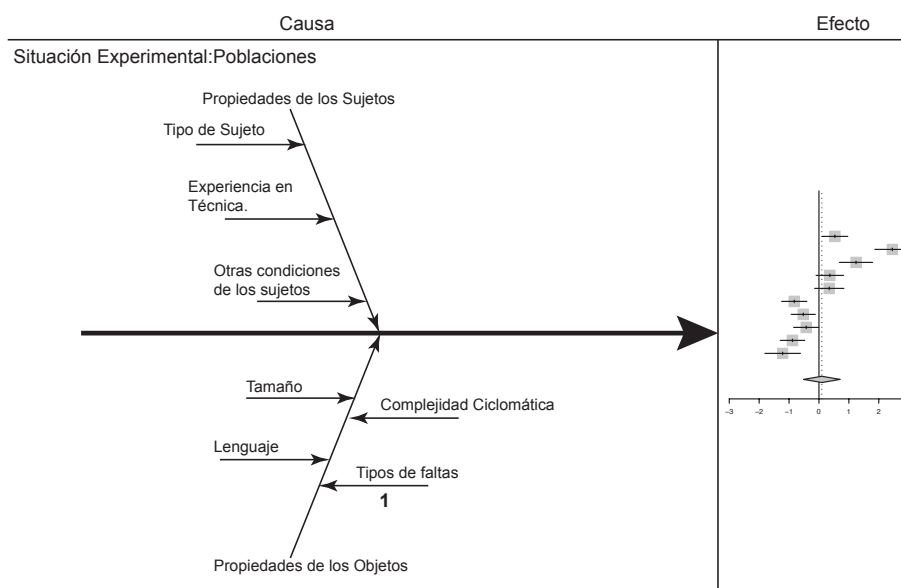


Fig. 12.22: Cambios identificados en poblaciones

contiene cierto número de faltas inyectadas en diferentes lugares del código fuente. De manera similar en las replicaciones *upm06* a *upm10* se usó una sola versión de los programas con faltas inyectadas en lugares distintos a las versiones usadas en las replicaciones *upm01* a *upm05*. Respecto al elemento de las operacionalizaciones que varía en los dos grupos se relaciona con las instrucciones que contienen el tratamiento a ser aplicado por los sujetos. La forma en como los sujetos siguen las instrucciones para aplicar las pruebas varía en las replicaciones de cada grupo de resultados.

Se cree que las faltas utilizadas en ambos grupos de replicaciones influyen en los resultados. Este cambio puede ser testado en futuras replicaciones para verificar esta hipótesis. La variación observada en las instrucciones de la prueba por cobertura de decisiones se cree que también influye en los resultados.

Una vez completada esta actividad del sub-proceso *análisis de resultados experimentales* se seleccionan los dos grupos de resultados consistentes para conocer el grado de verificación alcanzado y empaquetar este conocimiento tanto para el profesional como para el investigador. A continuación se desarrolla la última actividad de este sub-proceso.

12.2.3 Generar Tabla con Información del Análisis Realizado

En esta actividad se coloca en el formulario PCR-9 (apéndice D.9) la información de las comparaciones de resultados realizadas en la actividad anterior. En este caso se usan dos formularios, uno por cada grupo de resultados consistentes. En las tablas 12.6 y 12.7 se muestra la estructura principal del formulario PCR-9 con la información de los resultados de las comparaciones realizadas en la actividad anterior. Los grupos de resultados consistentes se han identificado como particiones 06.b.1 y 06.b.2.

12.3 Sub-proceso Empaquetado de Piezas de Conocimiento

Este es el último sub-proceso del *comparador de replicaciones* donde se empaquetan las piezas de conocimiento. Este sub-proceso se compone de las siguientes actividades:

- Generar tabla de variaciones en situación experimental
- Generar pieza de conocimiento para profesional
- Generar pieza de conocimiento para investigador

Información de la Partición	
Id. Partición:	06.b.1
Partición:	Efectividad/ Faltas Observadas/ Estruc. vs Lec. Cód./ Cob. Dec. vs. Lec. Cód. por A.S.
Replicaciones de la partición:	10 (upm01..upm10)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes
Información de los resultados:	<p>Q=155.9, $p<0.0001$, $I^2=94.2\%$, $T^2= 0.9183$</p> <p>Cob. Dec.: 51.41% [18.81; 83.99]</p> <p>f. observadas: 57.59, 84.88, 70.27, 63.38, 63.39 35.51, 40.97, 28.75, 38, 31.33</p> <p>Lec. Cód. por A.S.: 49.74% [27.49; 71.99]</p> <p>f. observadas: 47.27, 34.06, 46.3, 55.55, 54.91 53.62, 52.78, 36.81, 57.48, 58.66</p>
Tipo de inconsistencia observada:	inconsistencias en gráfico de control y meta-análisis
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	3 (upm01, upm04, upm05)
Comparación del par de tratamientos:	Efectividad en Cob. Dec. > Lec. Cód. por A.S.
Tipo de efecto global:	Significativo
Tamaño del efecto global:	0.4219 [0.1573; 0.6864]
Análisis de Causas	
Nuevas hipótesis de trabajo:	<p>1. El rol de monitor se cree influye en los resultados</p> <p>2. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados</p> <p>3. Las instrucciones en la prueba por Cob. Dec. se cree influyen en los resultados</p>

Tabla 12.6: Información de las comparaciones de resultados del primer grupo de resultados

Información de la partición					
Id Partición:		06.b.1			
Partición:		Efectividad/ Faltas Observadas/ Estruct. vs Lec. Cód./ Cob. Dec. vs Lec. Cód. por A.S.			
Replicaciones consideradas para la pieza de conocimiento:		3 (upm01, upm04, upm05)			
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	0	1	0%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	0	1	0%	60%
Sitio		0	1	0%	0%
Protocolo experimental	Análisis	0	1	0%	
	Diseño	0	6	0%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	0%
Operacionalizaciones de los constructos	Tratamiento A	1	5	20%	
	Tratamiento B	1	5	20%	
	Efecto	0	2	0%	13%
Poblaciones	Sujetos	0	3	0%	
	Objetos experimentales	0	4	0%	0%

Tabla 12.8: Tabla de variaciones en situación experimental de replicaciones upm01, upm04, upm05

12.3.2 Generar Pieza de Conocimiento para Profesional

A partir de las tablas con información de los análisis realizados en ambos grupos de resultados (sección 12.2.3) y con las tablas de variaciones (sección 12.3.1) se empaquetan las piezas de conocimiento para el profesional. Para construir las piezas de conocimiento se usa el formulario PCR-11 (apéndice D.11). En las tablas 12.10 y 12.11 se muestra la estructura principal del formulario PCR-11 con la información de las dos piezas de conocimiento resultantes para el profesional.

De manera general la información que ofrece esta primera pieza de conocimiento (tabla 12.10) al profesional es la siguiente. De diez replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al seleccionar el primer grupo de resultados (*upm01..05*) y apartar las replicaciones *upm02* y *upm03* la heterogeneidad es mínima por lo que se seleccionan las replicaciones *upm01*, *upm04* y *upm05* para obtener el grado de verificación logrado. Los resultados de la comparación entre ambas pruebas indican que la prueba por cobertura de decisiones es significativamente más efectiva que la prueba de lectura de código por abstracciones sucesivas. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser real debido a que no se ha variado el protocolo experimental. Respecto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. La pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas en esta pieza, por último el poder de confirmación de los

Información de la partición					
Id Partición:	06.b.2				
Partición:	Efectividad/ Faltas Observadas/ Estruct. vs Lec. Cód./ Cob. Dec. vs. Lec. Cód. por A.S.				
Replicaciones consideradas para la pieza de conocimiento:	4 (upm06..upm09)				
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	1	1	100%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	1	1	100%	100%
Sitio		0	1	0%	0%
Protocolo experimental	Análisis	0	1	0%	
	Diseño	3	6	50%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	10%
Operacionalizaciones de los constructos	Tratamiento A	0	5	0%	
	Tratamiento B	0	5	0%	
	Efecto	1	2	50%	17%
Poblaciones	Sujetos	0	3	0%	
	Objetos experimentales	0	4	0%	0%

Tabla 12.9: Tabla de variaciones en situación experimental de replicaciones upm06 a upm09

resultados es del 15%.

Respecto a la segunda pieza de conocimiento (tabla 12.11), la información que ésta ofrece al profesional es la siguiente. De diez replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al seleccionar el segundo grupo de resultados (*upm06..10*) y apartar *upm10* la heterogeneidad es mínima por lo que se seleccionan las replicaciones *upm06* a *upm09* para obtener el grado de verificación logrado. Los resultados de la comparación entre ambas pruebas indican que la prueba por cobertura de decisiones es significativamente menos efectiva que la prueba de lectura de código por abstracciones sucesivas. Este resultado es objetivo independiente de los experimentadores y local dependiente del sitio. El grado de verificación logrado es el siguiente, el resultado observado puede no ser completamente real debido a que no se ha variado la mayor parte del protocolo experimental. Respecto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas para esta pieza y el poder de confirmación de los resultados es del 25%.

Información de la Partición		
Id Partición:	06.b.1	
Partición:	Efectividad/ Faltas Observadas/ Estruc. vs Lec. Cód./ Cob. Dec. vs. Lec. Cód. por A.S.	
Replicaciones de la partición:	10 (upm01..upm10)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	Inconsistencias observadas en gráfico de control y meta-análisis	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	3 (upm01, upm04, upm05)	
Comparación:	Efectividad Cob Dec. > Lec. Cód. por A.S.	
Tipo de efecto global:	Significativo	
Tamaño del efecto global:	0.4219 [0.1573; 0.6864]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	60%	Semi-objetivo
Sitio:	0%	Local
Protocolo experimental:	0%	Artificial
Operacionalizaciones de los constructos:	13%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	15%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 12.10: Pieza de conocimiento para profesional del primer grupo de resultados

12.3.3 Generar Pieza de Conocimiento para Investigador

De manera similar a la actividad anterior, en esta última actividad se usan las tablas con información de los análisis realizados (sección 12.2.3) y las tablas de variaciones (sección 12.3.1). A partir de estas tablas se empaquetan las piezas de conocimiento para el investigador. Para construir estas dos piezas se utiliza el formulario PCR-12 (apéndice D.12). En las tablas 12.12 y 12.13 se muestra la estructura principal del formulario PCR-12 con la información de las piezas de conocimiento resultantes para el investigador.

De acuerdo a la información de esta primera pieza de conocimiento (tabla 12.12), el investigador obtiene los siguientes resultados. De la partición seleccionada se observa una heterogeneidad en el meta-análisis debido a los dos grupos de resultados observados. Al seleccionar el primer grupo (replicaciones *upm01..upm05*) se observa una heterogeneidad de resultados ocasionada *upm02* y *upm03*. Al apartar estas dos replicaciones los resultados son consistentes entre sí y se puede conocer el grado de verificación alcanzado. De acuerdo a la comparación de tratamientos, la prueba por cobertura de decisiones es significativamente más efectiva que la prueba de lectura de código por abstracciones sucesivas. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser real debido a que no se ha variado el protocolo experimental. Respecto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar

Información de la Partición		
Id Partición:	06.b.2	
Partición:	Efectividad/ Faltas Observadas/ Estruc. vs Lec. Cód./ Cob. Dec. vs. Lec. Cód. por A.S.	
Replicaciones de la partición:	10 (upm01..upm10)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	Inconsistencias observadas en gráfico de control y meta-análisis	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	4 (upm06..upm09)	
Comparación:	Efectividad en Cob. Dec. < Lec. Cód. por A.S.	
Tipo de efecto global:	Significativo	
Tamaño del efecto global:	-0.6649 [-0.8853; -0.4446]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	100%	Objetivo
Sitio:	0%	Local
Protocolo experimental:	10%	Parcialmente Artificial
Operacionalizaciones de los constructos:	17%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	25%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 12.11: Pieza de conocimiento para profesional del segundo grupo de resultados

los tratamientos y medir los resultados. Los tipos de replicaciones identificados son: *Nativa*, *Interna* y *Operacional*. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 15%.

De acuerdo al análisis de causas realizado, el rol del monitor, las instancias de las faltas así como el tipo de falta y las instrucciones en la prueba por cobertura de decisiones se cree influyen en los resultados. Estas condiciones pueden ser testadas en replicaciones posteriores.

Respecto a la información de la segunda pieza de conocimiento (tabla 12.13), el investigador obtiene los siguientes resultados. De la partición seleccionada se observa una heterogeneidad en el meta-análisis debido a los dos grupos de resultados observados. Al seleccionar el segundo grupo (replicaciones *upm06..upm10*) se observa una heterogeneidad de resultados ocasionada por *upm10*. Al apartar esta replicación los resultados son consistentes entre sí y es posible determinar el grado de verificación alcanzado. De acuerdo a la comparación de tratamientos, la prueba por cobertura de decisiones es significativamente menos efectiva que la prueba de lectura de código por abstracciones sucesivas.

El grado de verificación logrado es el siguiente, este resultado es objetivo independiente de los experimentadores y local dependiente del sitio. El resultado observado puede no ser completamente real debido a que no se ha variado la mayor parte del protocolo experimental. Respecto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados.

Información de la Partición			
Id Partición:	06.b.1		
Partición:	Efectividad/ Faltas Observadas/ Estruc. vs Lec. Cód./ Cob. Dec. vs. Lec. Cód. por A.S.		
Replicaciones de la partición:	10 (upm01..upm10)		
Resultados del análisis de similitudes:	Con Cambios		
Tipo de resultados encontrados:	Inconsistentes		
Información de los resultados:	Q=155.9, p<0.0001, I²=94.2%, T²= 0.9183 Cob. Dec.: 51.41% [18.81; 83.99] f. observadas: 57.59, 84.88, 70.27, 63.38, 63.39 35.51, 40.97, 28.75, 38, 31.33 Lec. Cód. por A.S.: 49.74% [27.49; 71.99] f. observadas: 47.27, 34.06, 46.3, 55.55, 54.91 53.62, 52.78, 36.81, 57.48, 58.66		
Tipo de inconsistencia observada:	inconsistencias en gráfico de control y meta-análisis		
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	3 (upm01, upm04, upm05)		
Comparación:	Efectividad en Cob. Dec. > Lec. Cód. por A.S.		
Tipo de efecto global:	Significativo		
Tamaño del efecto Global:	0.4219 [0.1573; 0.6864]		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	60%	Semi-objetivo	Nativa
Sitio:	0%	Local	Interna
Protocolo experimental:	0%	Artificial	-
Operacionalizaciones de los constructos:	13%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados	Operacional
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes	-
Poder de confirmación de resultados	15%		
Confiabilidad de la pieza de conocimiento:	Confiable		
Análisis de Causas			
Nuevas hipótesis de trabajo:	1. El rol de monitor se cree influye en los resultados 2. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados 3. Las instrucciones en la prueba por Cob. Dec. se cree influyen en los resultados		

Tabla 12.12: Pieza de conocimiento para investigador del primer grupo de resultados

Los tipos de replications identificados son: *Nativa*, *Interna*, *Artificial* y *Operacional*. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 25%.

De acuerdo al análisis de causas realizado, el tipo de sujeto, las instancias de las faltas así como el tipo de falta y las instrucciones en la prueba por cobertura de decisiones se cree influyen en los resultados. Estas condiciones pueden ser testadas en replicaciones futuras.

La tipología de replicaciones propuesta ha facilitado la identificación de replicaciones hechas de forma diferenciada, nativa e interna elegidas para realizar comparaciones y obtener conocimiento de ellas. Aunque se han observado inconsistencias en los resultados de las replicaciones de esta partición, a través del *comparador de replicaciones* ha sido posible conocer el grado de verificación alcanzado en las replicaciones con resultados consistentes. Este sub-proceso también ha facilitado la identificación de variables moderadoras que se cree influyen en los resultados experimentales. Estas variables pueden ser

Información de la Partición			
Id Partición:	06.b.2		
Partición:	Efectividad/ Faltas Observadas/ Estruc. vs Lec. Cód./ Cob. Dec. vs. Lec. Cód. por A.S.		
Replicaciones de la partición:	10 (upm01..upm10)		
Resultados del análisis de similitudes:	Con Cambios		
Tipo de resultados encontrados:	Inconsistentes		
Información de los resultados:	Q=155.9, p<0.0001, I ² =94.2%, T ² = 0.9183 Cob. Dec.: 51.41% [18.81; 83.99] f. observadas: 57.59, 84.88, 70.27, 63.38, 63.39 35.51, 40.97, 28.75, 38, 31.33 Lec. Cód. por A.S.: 49.74% [27.49; 71.99] f. observadas: 47.27, 34.06, 46.3, 55.55, 54.91 53.62, 52.78, 36.81, 57.48, 58.66		
Tipo de inconsistencia observada:	inconsistencias en gráfico de control y meta-análisis		
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	4 (upm06..upm09)		
Comparación:	Efectividad en Cob. Dec. < Lec. Cód. por A.S.		
Tipo de efecto global:	Significativo		
Tamaño del efecto Global:	-0.6649 [-0.8853; -0.4446]		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	100%	Objetivo	Nativa
Sitio:	0%	Local	Interna
Protocolo experimental:	10%	Parcialmente Artificial	Artificial
Operacionalizaciones de los constructos:	17%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados	Operacional
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes	-
Poder de confirmación de resultados	25%		
Confiabilidad de la pieza de conocimiento:	Confiable		
Análisis de Causas			
Nuevas hipótesis de trabajo:	1. El tipo de sujeto se cree influye en los resultados 2. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados 3. Las instrucciones en la prueba por Cob. Dec. se cree influyen en los resultados		

estudiadas en profundidad en replications posteriores.

Nótese que el estado de la práctica para llevar a cabo comparaciones entre replications era que sólo se comparaban los resultados experimentales y se intentaba agregarlos a través de técnicas de meta-análisis, en caso de obtener resultados no reproducibles, se fracasaba en la obtención de consecuencias de la comparación de los resultados. El *comparador de replications* permite al investigador comparar las situaciones experimentales de varias replications, comparar los resultados experimentales no sólo a través del meta-análisis sino también empleando otras técnicas que facilitan examinar las variaciones de los tratamientos. Éste también facilita la identificación de variables moderadoras cuando se obtienen resultados no reproducibles. Las distintas comparaciones hechas en este subproceso permiten conocer el grado de verificación logrado en las replications de una partición y empaquetar este conocimiento para que sea usado por el profesional o el investigador.

Capítulo 13

Evaluación con Replicaciones Diferenciadas-Nativas-Internas. Métrica Faltas Observables

RESUMEN: En este capítulo se evalúan las ideas del marco para la comparación de replicaciones referentes a las comparaciones entre replicaciones. Esta evaluación se realiza a través del sub-proceso *comparador de replicaciones* usando una partición compuesta por replicaciones realizadas de forma diferenciada, nativa e interna donde se utiliza otra métrica.

En los dos capítulos anteriores se evalúa el *comparador de replicaciones* con las particiones referentes a los pares de pruebas PCE vs. cobertura de decisiones así como cobertura de decisiones vs. lectura de código por abstracciones sucesivas. En este capítulo se analiza la partición 11.c (sección 10.2.7) que se compone de nueve replicaciones internas (*upm01..09*) realizadas por el grupo de investigación GrISE de la FI-UPM. En estas replicaciones se han variado principalmente elementos de la situación experimental referentes a los experimentadores, protocolo experimental y poblaciones. A diferencia de las dos particiones anteriores, en esta partición se compara la efectividad medida como el porcentaje de faltas observables¹ de las pruebas PCE y lectura de código por abstracciones sucesivas.

Al contar con la tabla maestra generada en 11.1, en este capítulo se omite esta actividad. A continuación se realizan los tres sub-procesos restantes (*identificación de similitudes*, *análisis de resultados experimentales* y *empaquetado de piezas de conocimiento*) del *comparador de replicaciones* usando la partición 11.c.

¹Que son las faltas observables en los casos de prueba especificados por el sujeto. Las faltas observables en un caso de prueba pueden no ser observadas por el sujeto. En el caso de la prueba por lectura de código, las faltas observables son las faltas que el sujeto ha descrito a través de las abstracciones, estas faltas pueden o no ser vistas por el sujeto.

13.1 Sub-proceso Identificación de Similitudes

Este sub-proceso se compone de las actividades *generar tablas de particiones* e *identificar similitudes*. A continuación se lleva a cabo cada una de estas actividades.

13.1.1 Generar Tablas de Particiones

Para generar las tablas de particiones se usa la información de la tabla maestra referente a las replicas *upm01..09* que conforman la partición 11.c.

En la tabla 13.1 se muestra la tabla de partición referente a los experimentadores y sitio. Las tablas completas de esta partición se encuentran en el apéndice G².

13.1.2 Identificar Similitudes

En esta actividad del sub-proceso se utilizan las tablas de particiones antes descritas más el procedimiento de automatización (apéndice B) para generar las representaciones de similitudes de esta partición.

Siguiendo la automatización de cálculos, primeramente se organizan las tablas de particiones de las diferentes secciones de la situación experimental en un fichero de texto CSV³ que representa la partición 11.c.

El fichero CSV de la partición 11.c se ingresa como parámetro al programa *Transpose* para generar ficheros de texto que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición. Los ficheros de texto resultantes se utilizan para generar de forma automática las representaciones de las similitudes. Este paso se realiza tras cargar el fichero `r_script01.txt` en el entorno R (R Dev. Core Team, 2011). Este fichero contiene varias funciones en R (R Dev. Core Team, 2011) para calcular las diferentes matrices de distancias así como las funciones para realizar los análisis de conglomerados jerárquico y de escalado multidimensional.

Una vez cargado en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`, se obtienen dos ficheros EPS que representan de forma gráfica las similitudes entre las replicas de la partición. En la figura 13.1 se muestran los gráficos resultantes que representan las similitudes de situaciones experimentales en esta partición⁴.

13.2 Sub-proceso Análisis de Resultados Experimentales

Este sub-proceso compara los resultados experimentales de las replicas de la partición y junto con los cambios identificados en las situaciones experimentales se determina el grado de verificación logrado en la partición. El sub-proceso *análisis de resultados experimentales* se compone de las siguientes actividades:

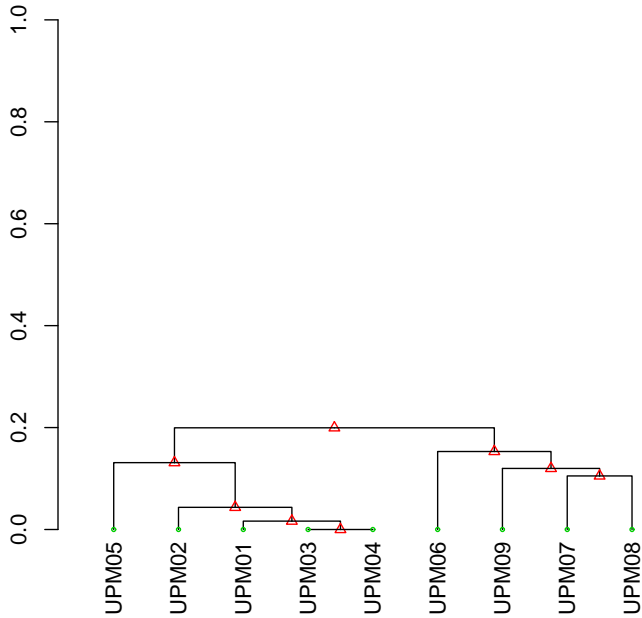
²En la tabla de particiones referente a las operacionalizaciones, el elemento identificado como *tratamiento como instrucciones* es igual para todas las replicas de la partición cuando se miden las faltas observables. Para obtener esta métrica, como primera actividad el sujeto debe generar casos de prueba (en el caso de las pruebas dinámicas) o abstracciones sucesivas (en el caso de la prueba por lectura de código). A partir de esta información se calcula el porcentaje de faltas observables por sujeto, por lo que el resto de las instrucciones del tratamiento quedan sin efecto para esta métrica.

³En el primer enlace de la sección *resources* de la url <http://www.osgg.net> se encuentra disponible el fichero de texto CSV de esta partición así como los ficheros generados de forma automática que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición.

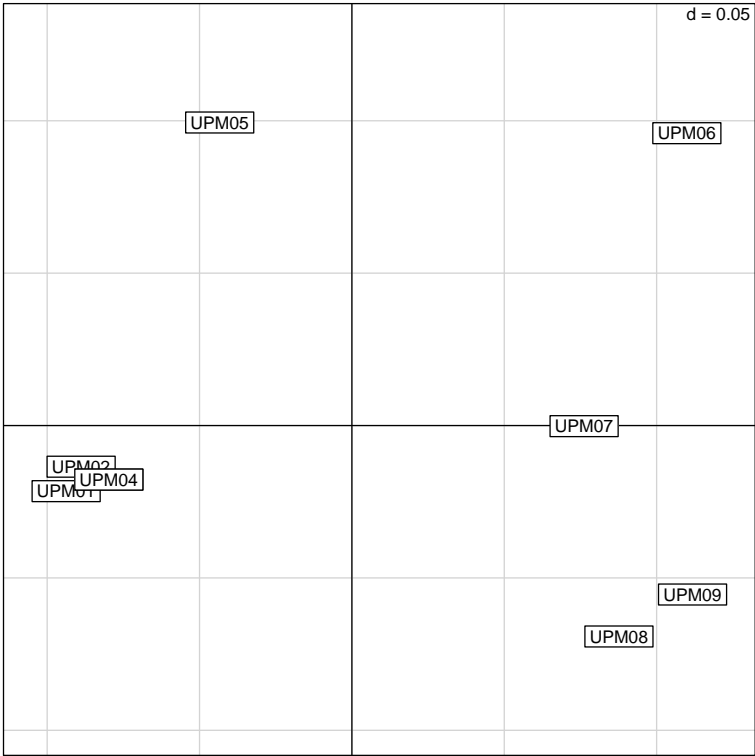
⁴Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`.

Estructura	Elemento principal	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09
Experimentadores	Diseñador	1.1	Vegas	X	X	X	X	X	X	X	X	
		1.1	Gómez									X
		1.1	Solari						X	X	X	
	Entrenador	1.2	Juristo	X	X	X	X	X	X	X	X	X
		1.2	Vegas	X	X	X	X			X	X	X
		1.2	Moreno					X	X			
	Monitor	1.3	Juristo	X	X	X	X			X	X	X
		1.3	Vegas	X	X	X	X	X	X	X	X	X
		1.3	Dieste		X							
		1.3	Ferre		X							
		1.3	Moreno					X	X			
		1.3	Gómez								X	X
	Medición	1.4	Vegas	X	X	X	X					
		1.4	Gómez								X	
		1.4	Solari					X	X	X		X
	Analista	1.5	Vegas	X	X	X	X	X				
		1.5	Gómez								X	X
		1.5	Solari						X	X		
Sitio		2	UPM	X	X	X	X	X	X	X	X	X

Tabla 13.1: Tabla de particiones referente a los experimentadores y sitio



(a) Dendrograma



(b) MDS

Fig. 13.1: Representación de similitudes, partición 11.c

- Generar tabla de resultados experimentales
- Realizar comparación de resultados experimentales
- Generar tabla con información del análisis realizado

13.2.1 Generar Tabla de Resultados Experimentales

En esta actividad se extraen los resultados experimentales de las replicaciones que conforman la partición 11.c y se colocan en el formulario PCR-7 (apéndice D.7). En la tabla 13.2 se presentan los resultados experimentales obtenidos en esta partición.

Id exp.	Tamaño (n_1)	Porcentaje faltas observables ($\bar{\gamma}_1$)	P.C.E.		Lec. Cód. por A.S.	Desv. est.
			Desv. est. (s_1)	Tamaño (n_2)	Porcentaje faltas observables ($\bar{\gamma}_2$)	Desv. est. (s_2)
upm01	42	82.48	22.08	42	47.28	18.8
upm02	39	81.01	17.75	39	35.53	20.43
upm03	29	82.35	25.4	29	46.31	18.96
upm04	36	82.54	16.87	36	55.56	23.6
upm05	31	82.57	14.82	32	54.91	24.1
upm06	46	47.1	17.67	46	53.62	23.02
upm07	48	42.36	15.36	48	52.78	21.01
upm08	48	40.63	19.72	48	36.81	18.5
upm09	48	47.22	15.88	49	57.48	19.26

Tabla 13.2: Resultados experimentales de la partición 11.c

13.2.2 Realizar Comparación de Resultados Experimentales

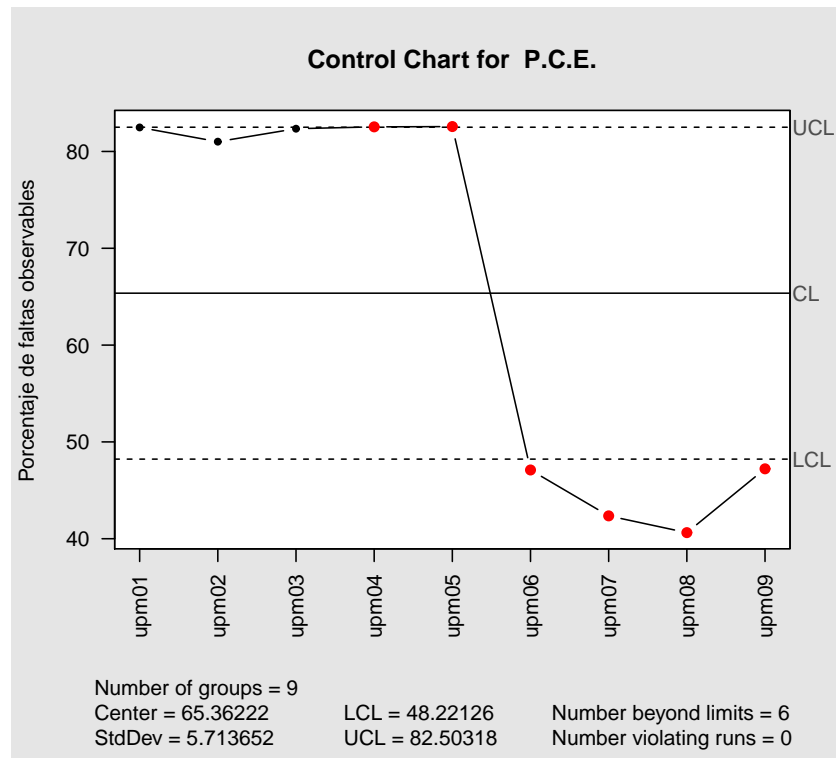
En esta actividad del sub-proceso se analizan los resultados experimentales de la partición. El análisis de los resultados se lleva a cabo a través del uso de gráficos de control y de técnicas estadísticas de agregación de resultados i.e. meta-análisis.

A partir de esta tabla se generan dos gráficos de control, uno por cada tratamiento así como el meta-análisis que contiene entre otras cosas los tamaños de efecto estimados de las replicaciones de la partición así como el tamaño de efecto global estimado. Estos análisis se llevan a cabo cargando el fichero `r_script02.txt` en el entorno R (R Dev. Core Team, 2011) para la partición 11.c. Este fichero recibe como entrada la tabla de resultados experimentales de la partición. Como salida se generan tres ficheros EPS, dos ficheros que contienen los gráficos de control del par de tratamientos (PCE y lec. de código por abs. sucesivas), y otro fichero que contiene los resultados del meta-análisis. Con los gráficos de control y meta-análisis resultantes más las representaciones de similitudes obtenidas previamente, a continuación se examinan en conjunto estos resultados.

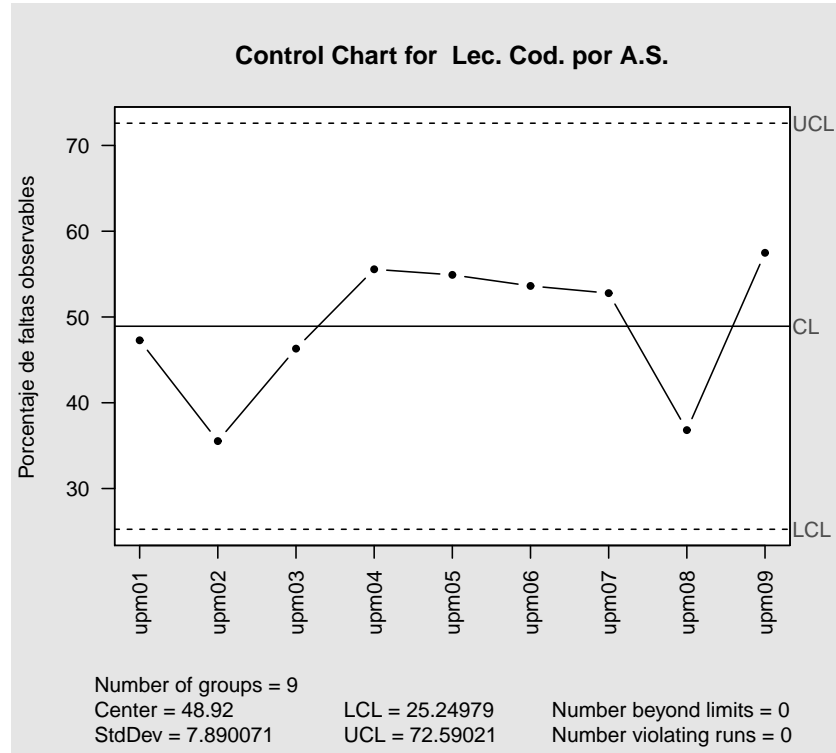
En la figura 13.2 se muestran los gráficos de control de los dos tratamientos examinados en esta partición (PCE y lec. de código por abs. sucesivas), mientras que en la figura 13.3 se muestran los resultados del meta-análisis⁵.

Como se observa en el gráfico de control de la prueba por PCE (figura 13.2a), los resultados de las replicaciones se encuentran fuera de los límites de control. En este gráfico se observan dos grupos de resultados. En el primer grupo de resultados (replicaciones *upm01* a *upm05*) se observa una efectividad de la prueba por PCE mayor al 80% mientras que en el segundo grupo de resultados (replicaciones *upm06* a *upm09*) se observa una efectividad menor al 50%. En este segundo grupo se observan resultados con más

⁵Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script02.txt`.



(a) Prueba por PCE



(b) Prueba por lec. de cód. por abs. suc.

Fig. 13.2: Gráficos de control resultantes

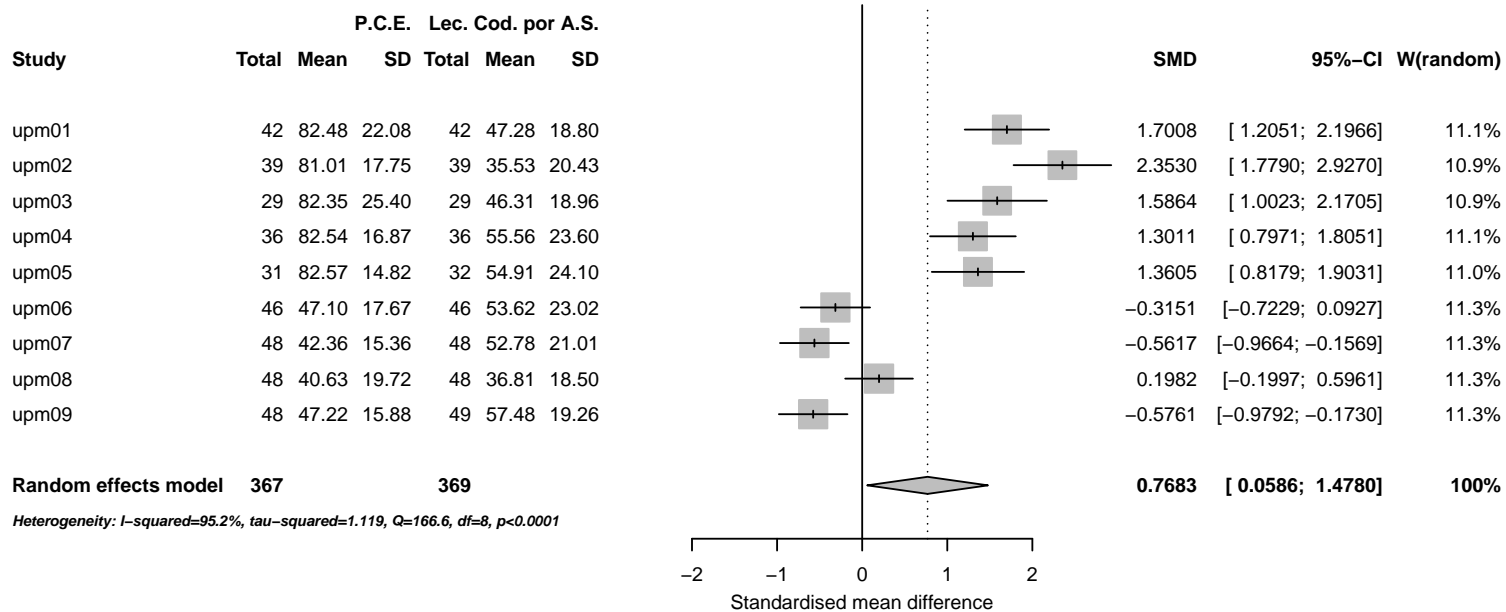


Fig. 13.3: Meta-análisis de las pruebas PCE y lec. de cód. por abs. sucesivas

variabilidad. Por otra parte, en la figura 13.2b que representa el gráfico de control de la prueba por lectura de código, los resultados de las replicaciones se encuentran dentro de los límites de control.

En la figura 13.3 se observan dos grupos de tamaños de efecto estimados, estos grupos ocasionan que la prueba de heterogeneidad sea significativa ($Q = 166.6$, $p < 0.0001$) de manera similar los valores de los índices I^2 y T^2 indican una heterogeneidad alta. El índice I^2 indica un porcentaje de variabilidad total en el conjunto de tamaños de efecto estimados del 95.2%, mientras que el índice T^2 que representa la varianza inter-estudios contiene un valor de 1.119.

Respecto al dendrograma (13.1a) y MDS 13.1b que representan las similitudes de las replicaciones, se observan dos agrupaciones, en un grupo se encuentran las replicaciones *upm01* a *upm05* mientras que en el otro las replicaciones *upm06* a *upm09*. En el caso del dendrograma la diferencia entre ambos grupos se produce a una distancia poco superior de 0.2 esto indica que la mayoría de los elementos de la situación experimental son comunes en la replicación de la partición.

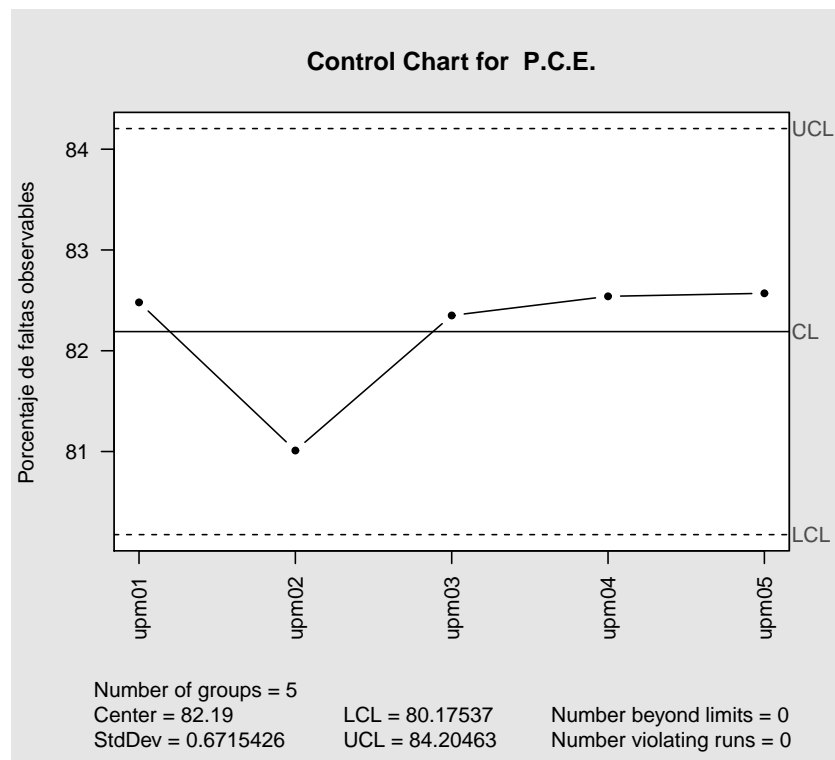
De acuerdo a los resultados presentados en los gráficos de control, meta-análisis e identificación de similitudes se tienen resultados inconsistentes producto del gráfico de control que representa la prueba por PCE y del meta-análisis, en los resultados de las representaciones de similitudes se observan cambios en la situación experimental. Si se emplean todas las replicaciones de esta partición para generar conocimiento, éste no sería confiable debido a las inconsistencias observadas. En esta situación, el investigador decide si intenta o no obtener conocimiento a partir del agrupamiento de resultados consistentes de la partición.

Con el fin de obtener algún conocimiento se decide avanzar e identificar grupos de resultados consistentes en esta partición. Siguiendo con esta actividad, a continuación se examina la distribución de los tamaños de efecto tanto en los gráficos de control como en el meta-análisis. En la distribución de resultados del gráfico de control de la prueba por PCE así como en la distribución de tamaños de efecto del meta-análisis se observan dos grupos de resultados. El primer grupo se compone de las replicaciones *upm01* a *upm05* mientras que el segundo grupo se compone de las replicaciones *upm06* a *upm09*. Con cada uno de estos grupos se vuelven a generar los gráficos de control y el meta-análisis para obtener resultados consistentes que permitan empaquetar el conocimiento obtenido en una o varias piezas. A continuación se examina cada uno de los grupos de resultados identificados.

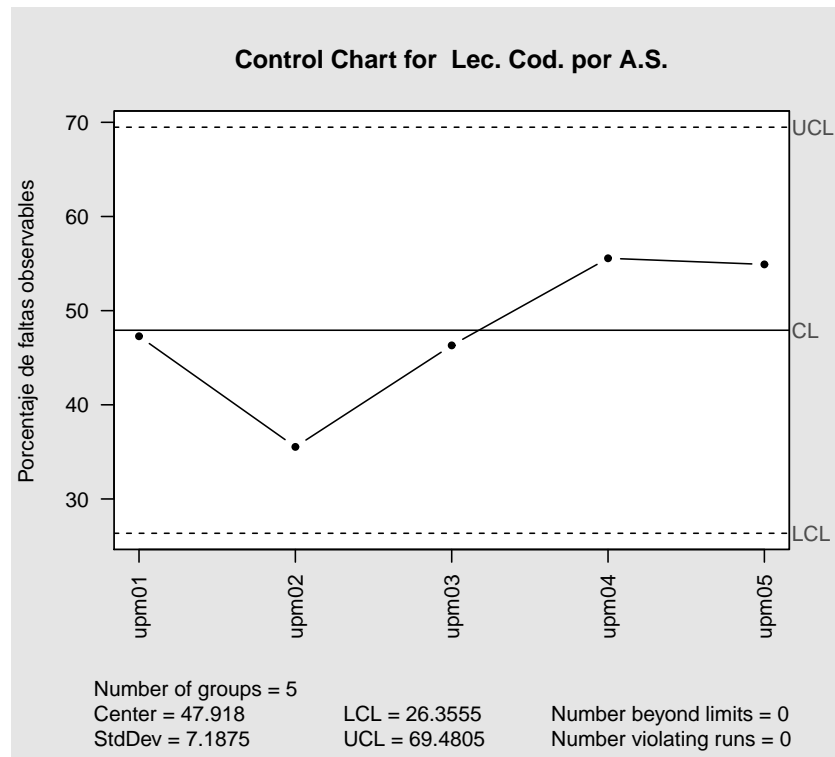
13.2.2.1 Análisis del Primer Grupo de Resultados

En las figuras 13.4 y 13.5 se muestran los gráficos de control y meta-análisis resultante del primer grupo de resultados,

En este primer grupo de resultados, los gráficos de control en ambos tratamientos (prueba por PCE y lec. de código por abstracciones sucesivas) se encuentran dentro de los límites de control, sin embargo en el meta-análisis se observa una heterogeneidad en los resultados debido en apariencia a *upm02*. En este grupo de replicaciones la prueba de heterogeneidad puede considerarse significativa con un valor $Q = 8.8$ y $p < 0.067$, el valor del índice I^2 indica un porcentaje de variabilidad del 54.4% i.e. variabilidad media, sin embargo el índice T^2 tiene un valor de 0.0898 que indica poca varianza inter-estudios. Ya que se observa una heterogeneidad media en los resultados se aparta *upm02* y se usa el resto de replicaciones de este grupo para obtener resultados consistentes. Con estas



(a) Prueba por PCE



(b) Prueba por lec. de cód. por abs. suc.

Fig. 13.4: Gráficos de control del primer grupo de resultados

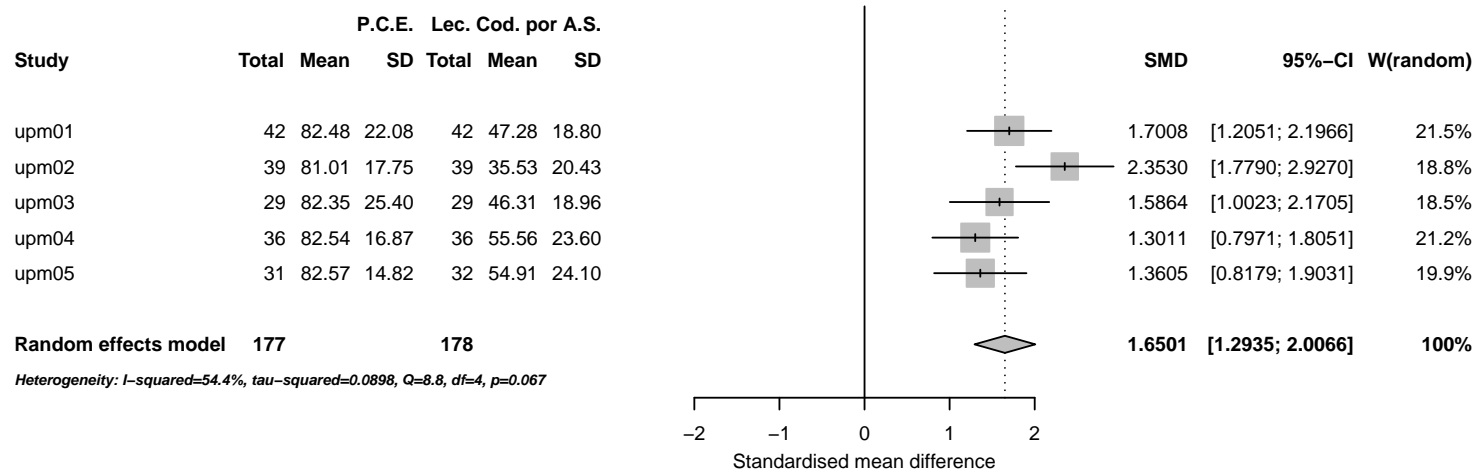


Fig. 13.5: Meta-análisis de las pruebas PCE y lec. de cód. por abs. sucesivas del primer grupo de resultados

cuatro replicaciones (*upm01*, *upm03*, *upm4* y *upm05*) se vuelven a generar los gráficos de control y meta-análisis. En las figuras 13.6 y 13.7 se muestran los resultados de los gráficos de control y meta-análisis tras apartar *upm02*.

Como se observa en las figuras 13.6a y 13.6b los resultados de ambas pruebas se encuentran dentro de los límites de control. Por otra parte, los resultados del meta-análisis (figura 13.7) son ya homogéneos. Ambos resultados tanto los gráficos de control como meta-análisis son ya consistentes y pueden usarse para empaquetar una pieza de conocimiento.

13.2.2.2 Análisis del Segundo Grupo de Resultados

Una vez que se identificaron resultados consistentes en el anterior grupo, a continuación se examina el segundo grupo de resultados compuesto por las replicaciones *upm06* a *upm09*. En las figuras 13.8 y 13.9 se muestran los gráficos de control y meta-análisis de este segundo grupo de resultados.

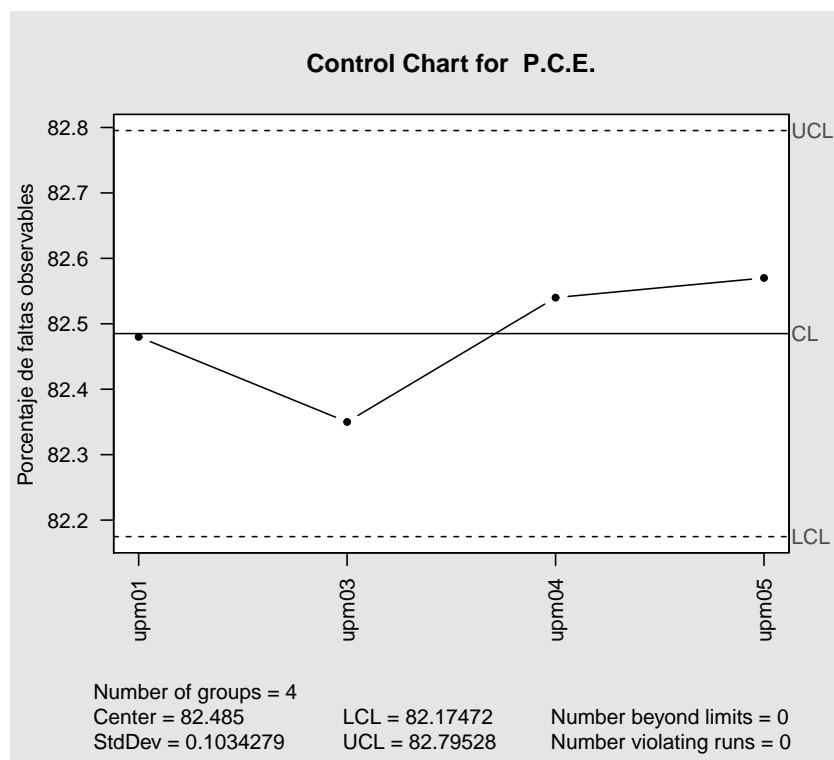
En este segundo grupo de resultados, los gráficos de control en ambos tratamientos (prueba por PCE y lec. de código por abstracciones sucesivas) se encuentran dentro de los límites de control, sin embargo en el meta-análisis se observa una heterogeneidad en los resultados debido en apariencia a la replicación *upm08*. En este grupo de replicaciones la prueba de heterogeneidad es significativa con un valor $Q = 9.4$ y $p < 0.0241$, el valor del índice I^2 indica un porcentaje de variabilidad del 68.2% y el índice T^2 un valor de 0.0907. Para disminuir la heterogeneidad en los resultados se aparta *upm08* y se usa el resto de replicaciones de este grupo para obtener resultados consistentes. Con las tres replicaciones restantes (*upm06*, *upm07* y *upm09*) se vuelven a generar los gráficos de control y meta-análisis. En las figuras 13.10 y 13.11 se muestran los resultados de los gráficos de control y meta-análisis después de apartar *upm08*.

Como se observa en las figuras 13.10a y 13.10b los resultados de ambas pruebas se encuentran dentro de los límites de control. Por otra parte, los resultados del meta-análisis (figura 13.11) son homogéneos. Ambos resultados tanto los gráficos de control como meta-análisis son consistentes y pueden usarse para conocer el grado de verificación y empaquetar este conocimiento, no obstante al apartar replicaciones en esta partición se recomienda llevar a cabo el análisis de causas para tratar de identificar los elementos de la situación experimental que se cree influyen en los resultados experimentales. A continuación se realiza el análisis de causa en cada uno de los grupos de resultados (intra-grupos) y por último a nivel de grupos de resultados (inter-grupos).

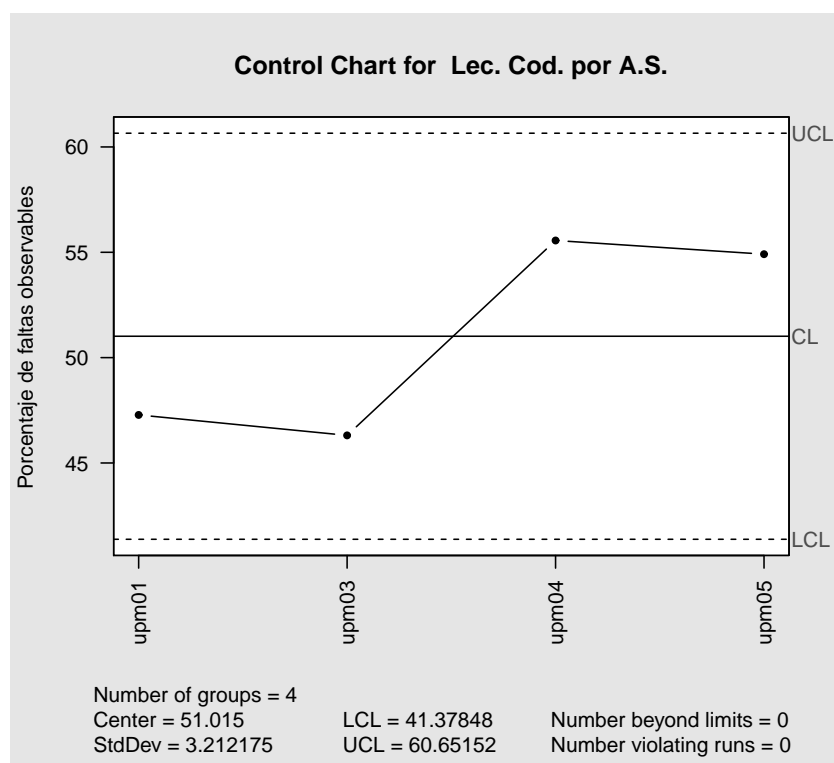
13.2.2.3 Análisis de Causas del Primer Grupo de Resultados

En las figuras 13.12 y 13.13 se muestran dos diagramas de Ishikawa (Ishikawa, 1990) con el número de cambios identificados en los elementos de la situación experimental del primer grupo de resultados (replicaciones *upm01* a *upm05*). En la tabla 13.3 se muestra el fragmento de las tablas de particiones con los cambios identificados.

Respecto a estos cambios, se identifica un cambio en el elemento *aspectos de transmisión del tratamiento* de las operacionalizaciones de los constructos. A diferencia del resto de replicaciones, en *upm01* participaron dos investigadores durante la sesión del entrenamiento uno de ellos con experiencia para transmitir el tratamiento mientras que el otro carecía de experiencia, en el resto de replicaciones, los investigadores cuentan con experiencia para transmitir el tratamiento a los sujetos. El cambio identificado en *upm01* parece no influir en los resultados, por ejemplo, se verifican los resultados entre *upm01*



(a) Prueba por PCE



(b) Prueba por lec. de cód. por abs. suc.

Fig. 13.6: Gráficos de control del primer grupo de resultados sin upm02

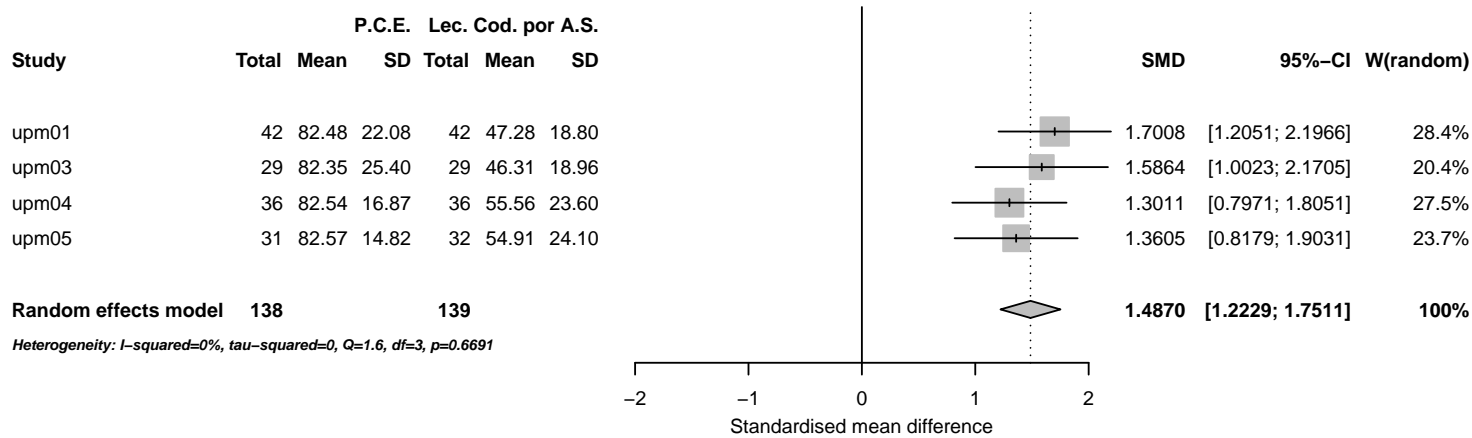
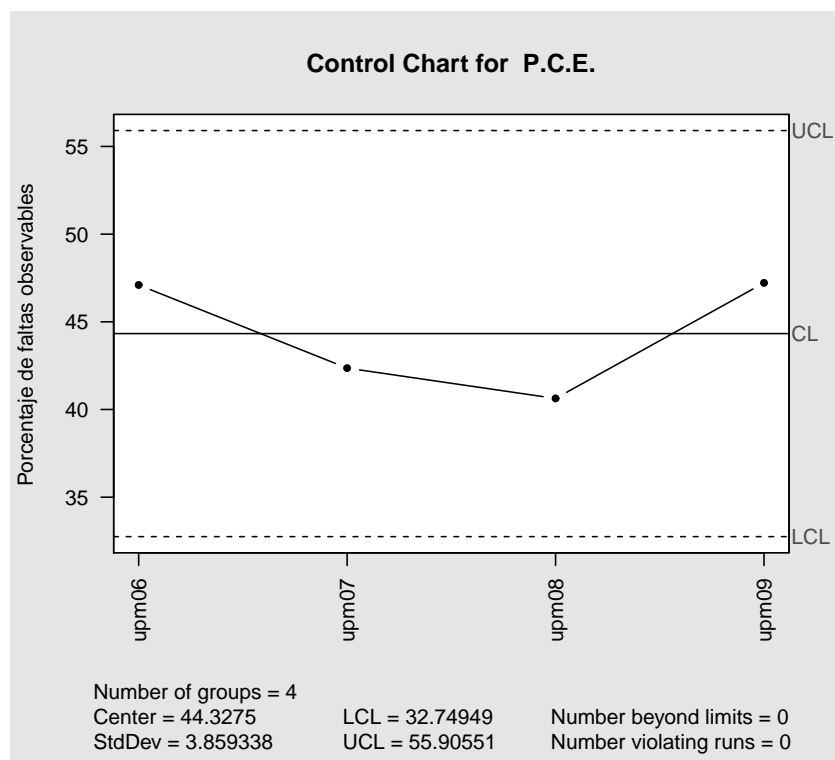
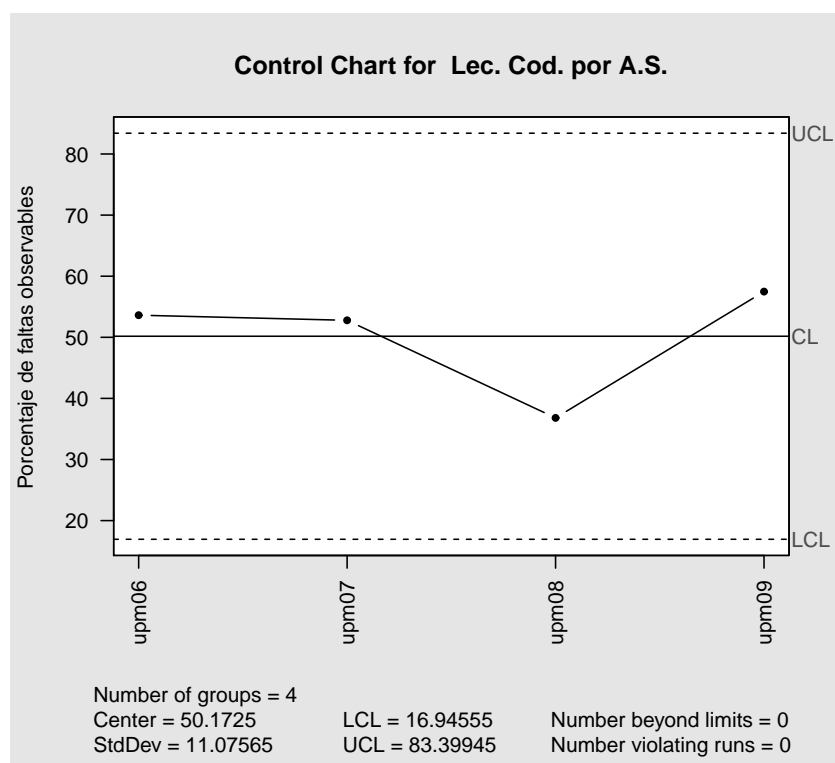


Fig. 13.7: Meta-análisis de las pruebas PCE y lec. de cód. por abs. sucesivas del primer grupo de resultados sin upm02



(a) Prueba por PCE



(b) Prueba por lec. de cód. por abs. suc.

Fig. 13.8: Gráficos de control del segundo grupo de resultados

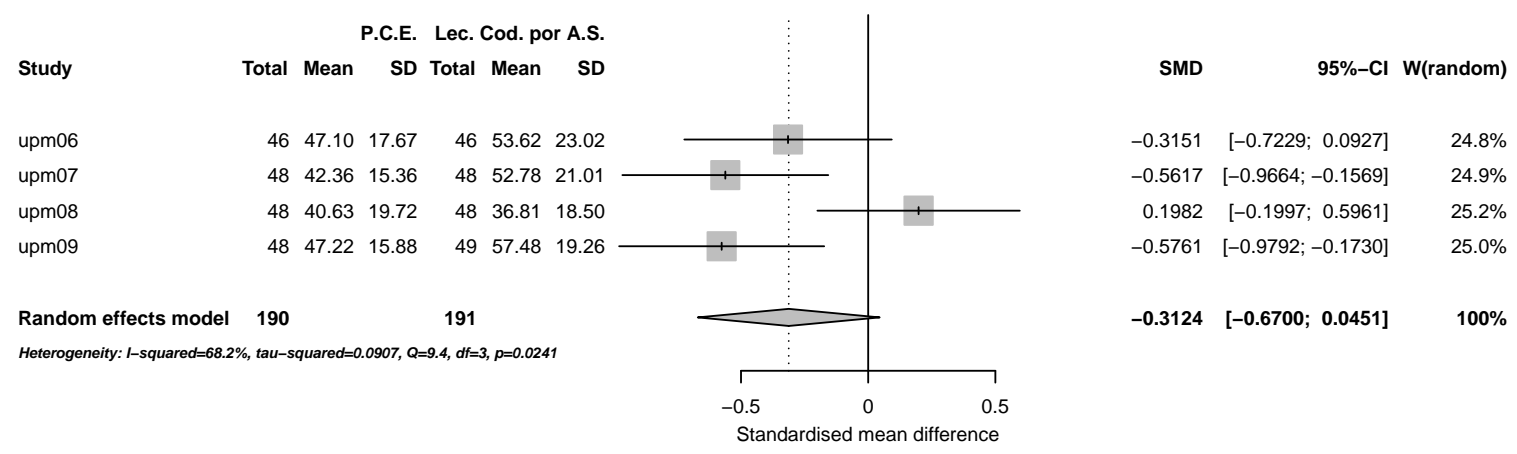
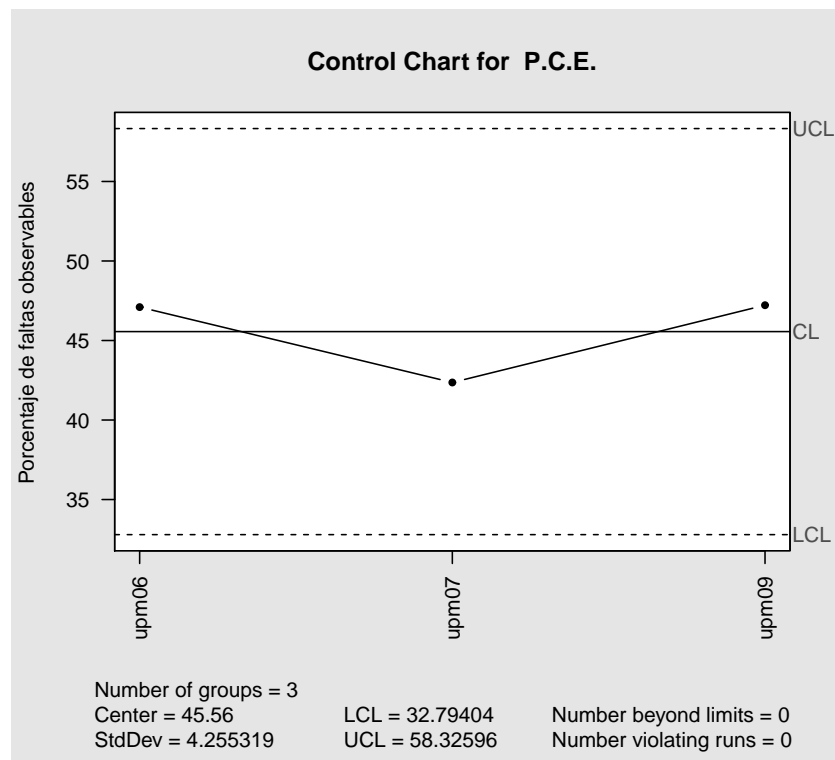
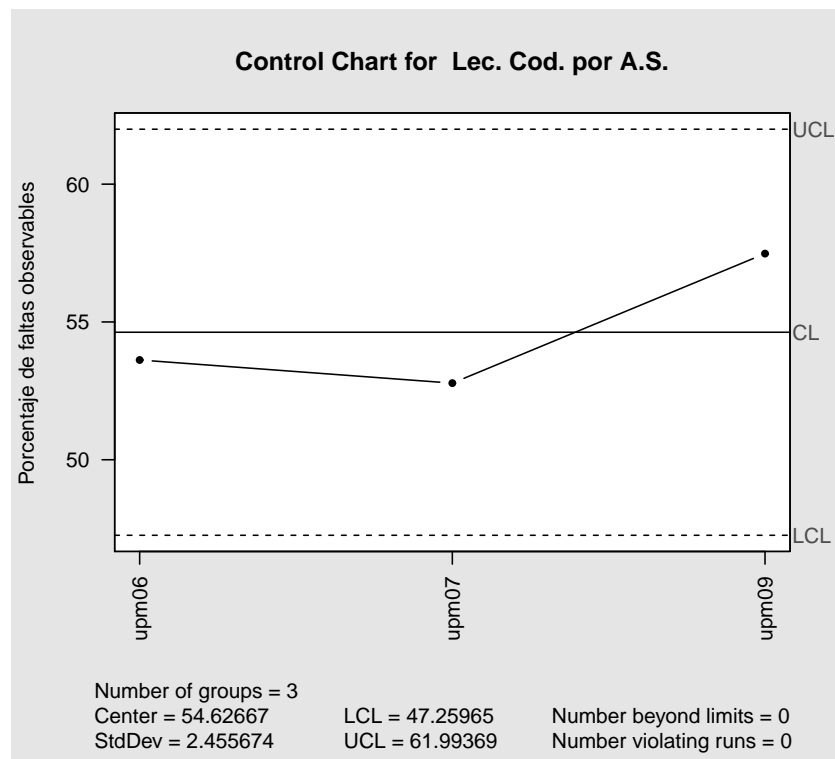


Fig. 13.9: Meta-análisis de las pruebas PCE y lec. de cód. por abs. sucesivas del segundo grupo de resultados



(a) Prueba por PCE



(b) Prueba por lec. de cód. por abs. suc.

Fig. 13.10: Gráficos de control del segundo grupo de resultados sin upm08

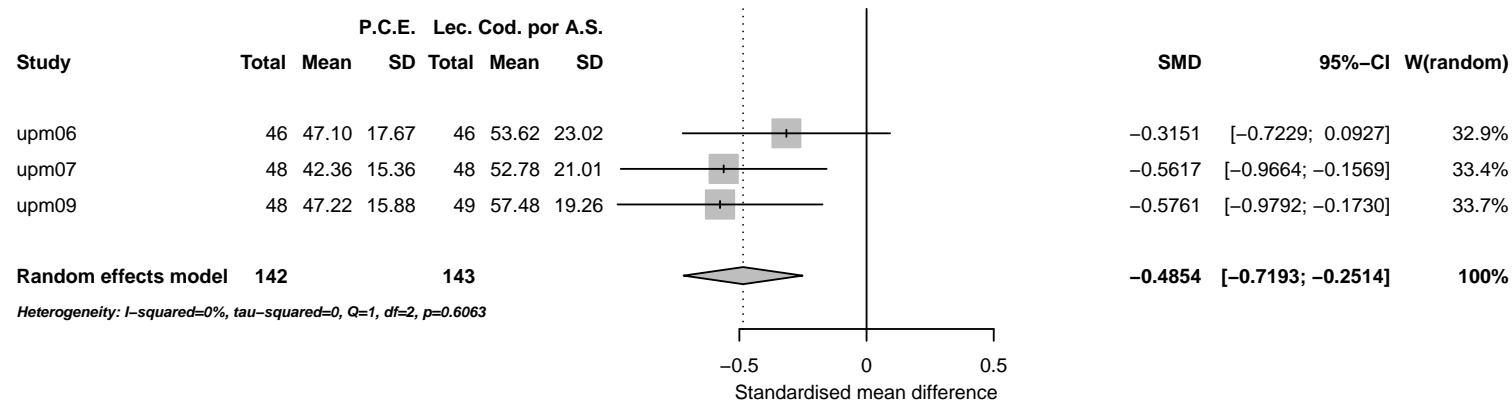


Fig. 13.11: Meta-análisis de las pruebas PCE y lec. de cód. por abs. sucesivas del segundo grupo de resultados sin upm08

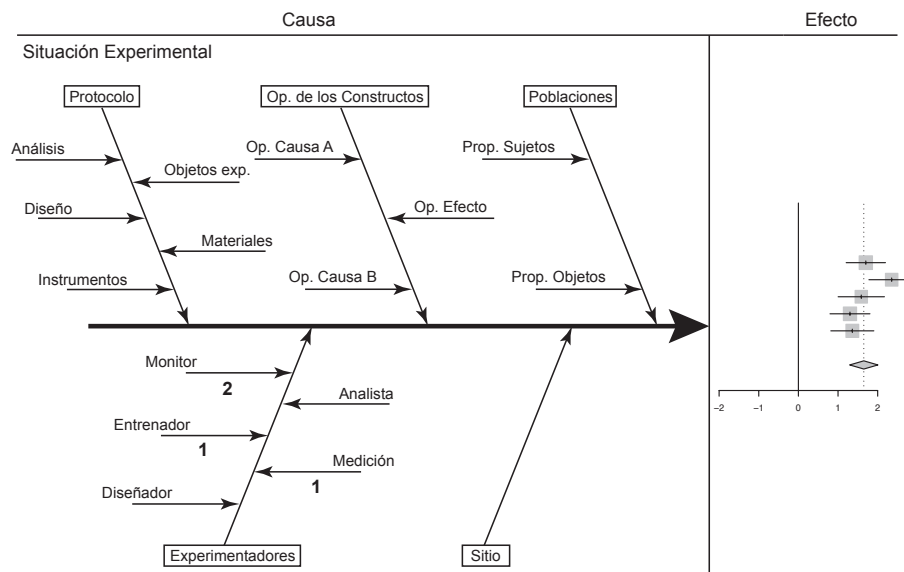


Fig. 13.12: Cambios identificados en experimentadores y sitio

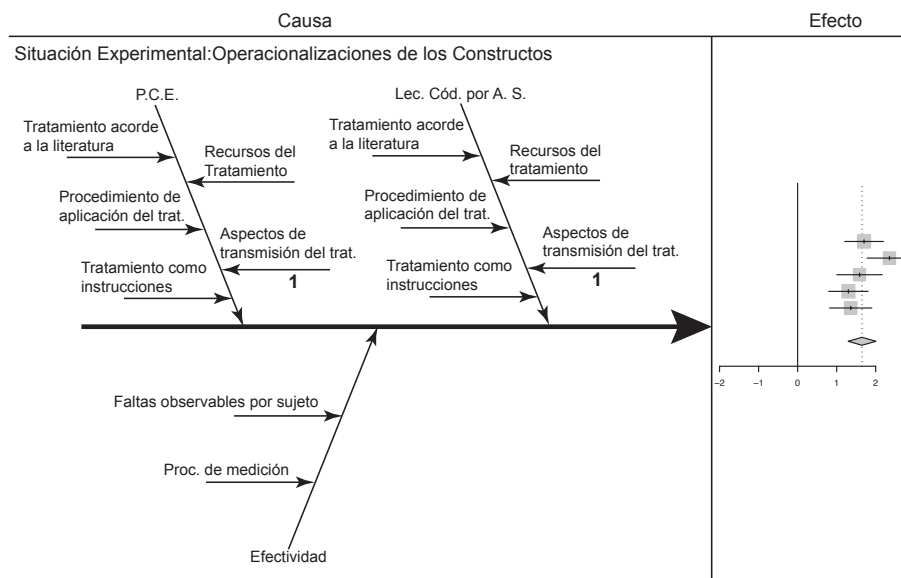


Fig. 13.13: Cambios identificados en operacionalizaciones de los constructos

Elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05
Entrenador	1.2	Juristo	X	X	X	X	X
	1.2	Vegas	X	X	X	X	
	1.2	Moreno					X
Monitor	1.3	Juristo	X	X	X	X	
	1.3	Vegas	X	X	X	X	X
	1.3	Dieste		X			
	1.3	Ferre		X			
	1.3	Moreno					X
Medición	1.4	Vegas	X	X	X	X	
	1.4	Solari					X
Cob. Dec.	Aspectos de transmisión del tratamiento	4.1.1.5 Con experiencia en curso universitario	X	X	X	X	X
		4.1.1.5 Sin experiencia en curso universitario	X				
Lec. Cód. A.S.	Aspectos de transmisión del tratamiento	4.1.2.5 Con experiencia en curso universitario	X	X	X	X	X
		4.1.2.5 Sin experiencia en curso universitario	X				

Tabla 13.3: Cambios identificados en los elementos de la situación experimental

y *upm04* (figura 12.5). En este par de replicaciones sólo se varía este elemento de la situación experimental.

En *upm05* se observan cambios en los roles de entrenador, monitor y medición respecto a las replicaciones *upm01*, *upm03* y *upm04* que utilizan los mismos roles. Los cambios en los roles en *upm05* parecen no influir ya que se verifican los resultados con las replicaciones *upm01* y *upm04*.

En el caso de *upm02* los cambios en el rol de monitor parecen influir en los resultados. Esta replicación es más similar a *upm03*. A diferencia de *upm03*, en *upm02* participan otros dos investigadores en el rol de monitores. Tal vez estos investigadores dieron algunas instrucciones a los sujetos que influyeron en los resultados de este experimento.

Por otra parte, la replicación *upm03* es igual a *upm04* sin embargo los resultados de la replicación *upm03* difieren del resto. En esta replicación otra condición no identificada pudo influir en los resultados.

De los cambios anteriormente discutidos se cree que el rol de monitor puede influir en los resultados. Este elemento puede ser testado en futuras replicaciones para verificar esta hipótesis.

13.2.2.4 Análisis de Causas del Segundo Grupo de Resultados

En las figuras 13.14, 13.15 y 13.16 se usan los diagramas de Ishikawa (Ishikawa, 1990) para representar los cambios identificados en las replicaciones (*upm06* a *upm09*) de este segundo grupo de resultados. En la tabla 13.4 se muestra la porción de las tablas de particiones con los cambios identificados.

En este segundo grupo, los resultados de *upm08* ocasionan un aumento en la heterogeneidad de los resultados (figura 12.11). A diferencia del resto de replicaciones, se observa un cambio en la dirección del tamaño de efecto estimado en *upm08*. El cambio notorio que hace diferente a la replicación *upm08* del resto es el investigador que participó en el rol de medición y el criterio de corrección usado como parte del procedimiento de medición. Una vez concluido el experimento, durante la revisión de las diferentes pruebas aplicadas por los sujetos, el revisor (rol de medición) usó un criterio estricto de revisión. Se cree que este criterio pudo influir en los resultados de la replicación. En esta replicación el porcentaje de efectividad en ambas pruebas es ligeramente menor que el resto, en los

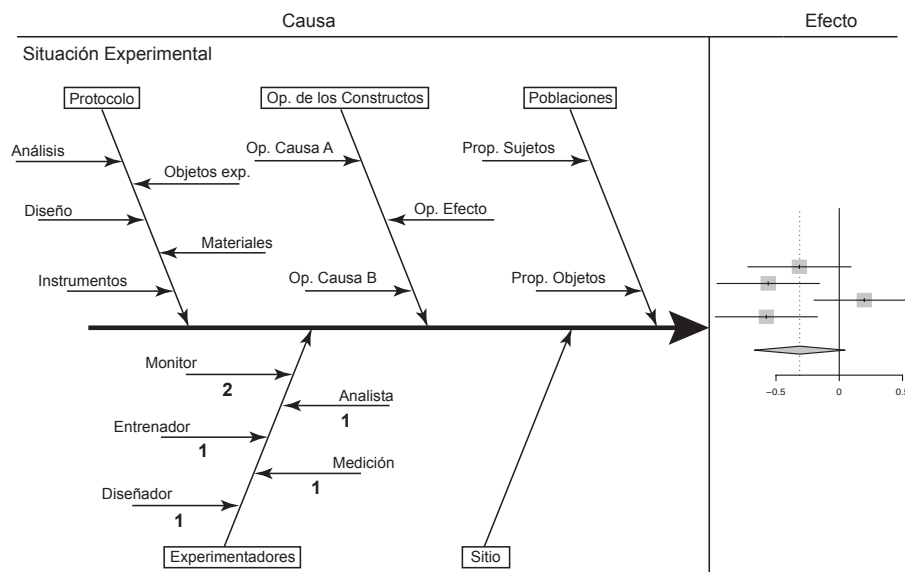


Fig. 13.14: Cambios identificados en experimentadores y sitio

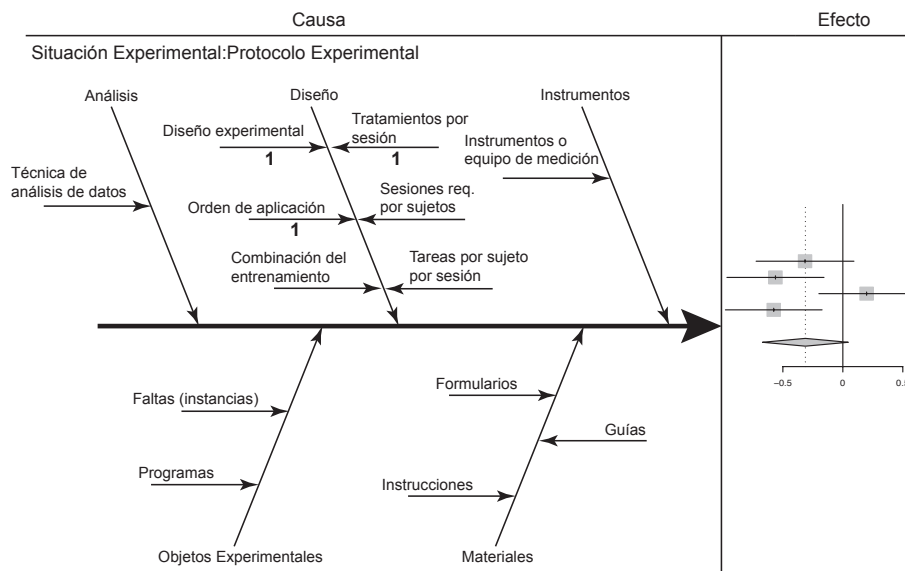


Fig. 13.15: Cambios identificados en el protocolo experimental

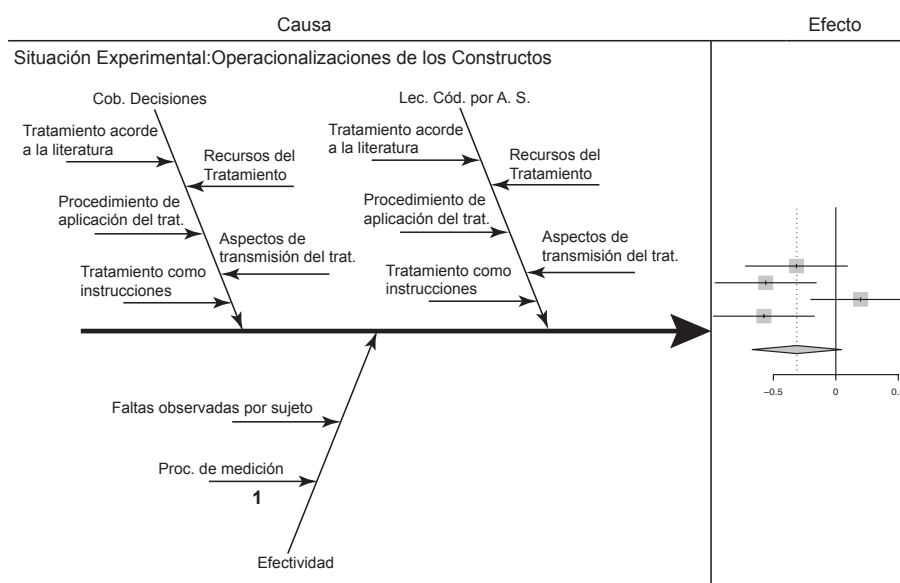


Fig. 13.16: Cambios identificados en operacionalizaciones de los constructos

Elemento	ID	Categoría	UPM06	UPM07	UPM08	UPM09
Diseñador	1.1	Vegas	X	X	X	
	1.1	Gómez				X
	1.1	Solari	X	X	X	
Entrenador	1.2	Juristo	X	X	X	X
	1.2	Vegas		X	X	X
	1.2	Moreno	X			
Monitor	1.3	Juristo		X	X	X
	1.3	Vegas	X	X	X	X
	1.3	Moreno	X			
	1.3	Gómez			X	X
Medición	1.4	Vegas				
	1.4	Gómez			X	
	1.4	Solari	X	X		X
Analista	1.5	Gómez			X	X
	1.5	Solari	X	X		
Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X	
	3.2.1	Diseño Factorial				X
Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X	
	3.2.2	No Aplica (c/ suj app 1 tec)				X
Tratamientos por sesión	3.2.4	3 Técnicas, 1 Programa	X	X	X	
	3.2.4	3 Técnicas, 3 Programas				X
Procedimiento de medición (criterio de corrección)	4.2.2	Criterio relajado	X	X		X
	4.2.2	Criterio estricto			X	

Tabla 13.4: Cambios identificados en los elementos de la situación experimental

gráficos de control 13.8a y 13.8b se observa este pequeño decremento de efectividad. Se cree que también esta situación se debe al modo de revisión que se usó en esta replicación. El resto de cambios observados en las replicaciones de este grupo parecen no influir en los resultados.

De los cambios identificados en este segundo grupo de resultados se cree que el rol de medición así como el procedimiento de medición pueden influir en los resultados. Estos dos elementos pueden ser testados en futuras replicaciones para verificar estas hipótesis iniciales.

13.2.2.5 Análisis de Causas Inter-grupos de Resultados

Una vez analizados los cambios a nivel intra-grupos, a continuación se realiza el análisis de causas inter-grupos. En las figuras 13.17 y 13.18 se muestran los diagramas de Ishikawa (Ishikawa, 1990) con el número de cambios identificados entre estos dos grupos de resultados.

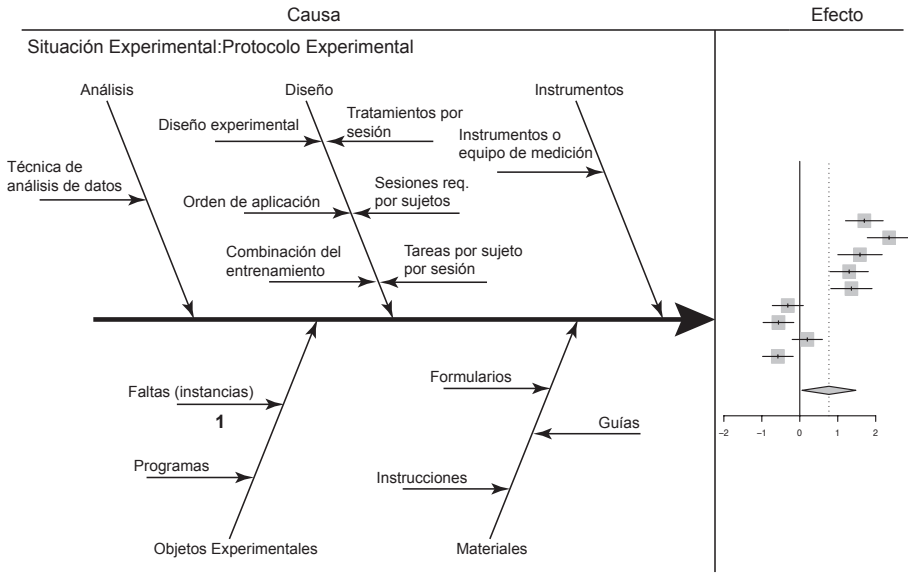


Fig. 13.17: Cambios identificados en el protocolo experimental

Elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09
Faltas (Instancias)	3.4.2	Instancia G	X	X	X	X	X				
	3.4.2	Instancia H	X	X	X	X	X				
	3.4.2	Instancia I						X	X	X	X
Tipos de Faltas	5.2.4	Grupo de Faltas G	X	X	X	X	X				
	5.2.4	Grupo de Faltas H	X	X	X	X	X				
	5.2.4	Grupo de Faltas I						X	X	X	X

Tabla 13.5: Cambios identificados en los elementos de la situación experimental

Los cambios identificados en los elementos de la situación experimental entre estos dos grupos de resultados se relacionan con las instancias de faltas de los objetos experimentales y con los tipos de faltas. En las replicaciones *upm01* a *upm05* se usaron dos versiones de cada programa. Cada versión tiene cierto número de faltas inyectadas en diferentes

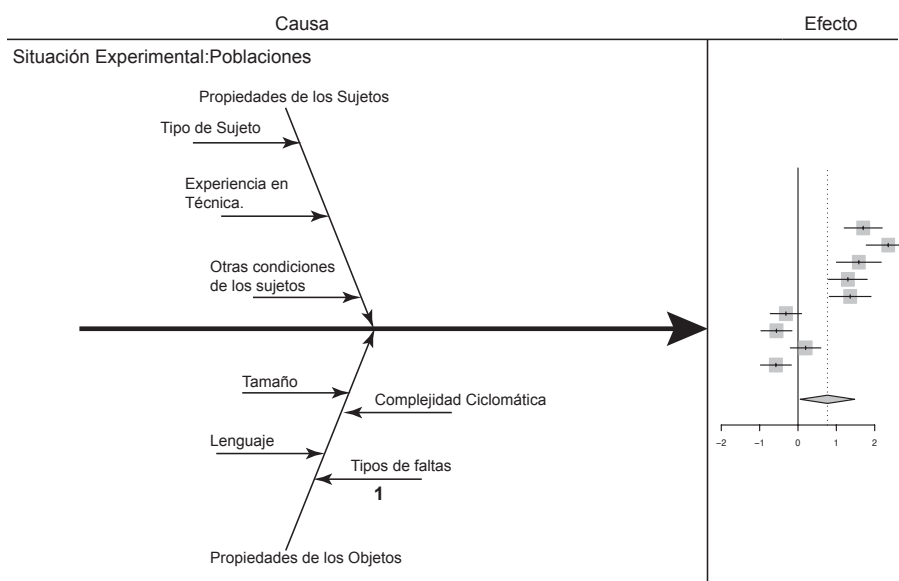


Fig. 13.18: Cambios identificados en poblaciones

lugares del código fuente. De manera similar en las replicaciones *upm06* a *upm09* se usó una sola versión de los programas con faltas inyectadas en lugares distintos a las versiones usadas en las replicaciones *upm01* a *upm05*. Se cree que las faltas utilizadas en ambos grupos de replicaciones influyen en los resultados. Este cambio puede ser testado en futuras replicaciones para verificar esta hipótesis.

En el capítulo anterior (sección 12.2.2.5), también se identificó que el elemento *tratamiento como instrucciones* pudo influir en los resultados, si este elemento es responsable de influir en los resultados de ambos grupos entonces en esta partición los resultados deberían ser consistentes entre las replicaciones de la partición, sin embargo, en esta partición también se observan dos grupos de resultados, por lo que la hipótesis respecto a la influencia de las faltas en los resultados toma mayor fuerza.

Una vez completada esta actividad del sub-proceso *análisis de resultados experimentales* se seleccionan los dos grupos de resultados consistentes para intentar empaquetar las piezas de conocimiento. A continuación se desarrolla la última actividad de este sub-proceso.

13.2.3 Generar Tabla con Información del Análisis Realizado

Una vez realizada la comparación de resultados experimentales, en esta actividad se coloca en el formulario PCR-9 (apéndice D.9) la información de las comparaciones de resultados realizadas en la actividad anterior, esta información se utiliza para determinar el grado de verificación alcanzado y empaquetar este conocimiento. En este caso se usan dos formularios, uno por cada grupo de resultados consistentes. En las tablas 13.6 y 13.7 se muestra la estructura principal del formulario PCR-9 con la información de los resultados de las comparaciones realizadas en la actividad anterior. Los grupos de resultados consistentes se han identificado como particiones 11.c.1 y 11.c.2.

Información de la Partición	
Id. Partición:	11.c.1
Partición:	Efectividad/ Faltas Observables/ Func. vs Lec. Cód./ P.C.E. vs. Lec. Cód. por A.S.
Replicaciones de la partición:	9 (upm01..upm09)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes
Información de los resultados:	Q=166.6, $p<0.0001$, $I^2=95.2\%$, $T^2= 1.119$ P.C.E.: 65.36% [48.22; 82.5031] f. observables: 82.48, 81.01, 82.35, 82.54, 82.57 47.1, 42.36, 40.63, 47.22 Lec. Cód. por A.S.: 48.92% [25.25; 72.59] f. observables: 47.28, 35.53, 46.31, 55.56, 54.91 53.62, 52.78, 36.81, 57.48
Tipo de inconsistencia observada:	inconsistencias en el meta-análisis y gráficos de control
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	4 (upm01, upm03, upm04, upm05)
Comparación del par de tratamientos:	P.C.E. > Lec. Cód. por A.S.
Tipo de efecto global:	Significativo
Tamaño del efecto global:	1.487 [1.2229; 1.7511]
Análisis de Causas	
Nuevas hipótesis de trabajo:	1. El rol de monitor se cree influye en los resultados 2. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados

Tabla 13.6: Información de las comparaciones de resultados del primer grupo de resultados

13.3 Sub-proceso Empaquetado de Piezas de Conocimiento

Una vez realizadas las actividades de los dos sub-procesos anteriores, a continuación se llevan a cabo las actividades del último sub-proceso del *comparador de replicaciones* para empaquetar las piezas de conocimiento. Este sub-proceso se compone de las siguientes actividades:

- Generar tabla de variaciones en situación experimental
- Generar pieza de conocimiento para profesional
- Generar pieza de conocimiento para investigador

Información de la Partición	
Id. Partición:	11.c.2
Partición:	Efectividad/ Faltas Observables/ Func. vs Lec. Cód./ P.C.E. vs. Lec. Cód. por A.S.
Replicaciones de la partición:	9 (upm01..upm09)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes
Información de los resultados:	Q=166.6. p<0.0001. I ² =95.2%. T ² = 1.119 P.C.E.: 65.36% [48.22; 82.5] f. observables: 82.48, 81.01, 82.35, 82.54, 82.57 47.1, 42.36, 40.63, 47.22 Lec. Cód. por A.S.: 48.92% [25.25; 72.59] f. observables: 47.28, 35.53, 46.31, 55.56, 54.91 53.62, 52.78, 36.81, 57.48
Tipo de inconsistencia observada:	inconsistencias en el meta-análisis y gráficos de control
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	3 (upm06, upm07, upm09)
Comparación del par de tratamientos:	P.C.E. < Lec. Cód. por A.S.
Tipo de efecto global:	Significativo
Tamaño del efecto global:	-0.4854 [-0.7193; -0.2514]
Análisis de Causas	
Nuevas hipótesis de trabajo:	1. El rol de medición se cree puede influir en los resultados 2. El procedimiento de medición se cree influye en los resultados 3. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados

Tabla 13.7: Información de las comparaciones de resultados del segundo grupo de resultados

13.3.1 Generar Tabla de Variaciones en Situación Experimental

En esta actividad se utiliza la información de las tablas de particiones generadas en el sub-proceso *identificación de similitudes* para calcular los cambios realizados en la situación experimental de las replicaciones que constituyen los resultados consistentes, en este caso las replicaciones del primer grupo de resultados *upm01*, *upm03*, *upm04* y *upm05* y las replicaciones del segundo grupo de resultados *upm06*, *upm07* y *upm09*. En esta actividad se usa el formulario PCR-10 (apéndice D.10) para colocar los cálculos de los cambios realizados. La tabla resultante se utiliza en las siguientes actividades del sub-proceso para empaquetar las piezas de conocimiento. En las tablas 13.8 y 13.9 se muestra la estructura principal del formulario PCR-10 con la el cálculo de variaciones en ambos grupos de resultados consistentes.

Información de la partición					
Id Partición:		11.c.1			
Partición:		Efectividad/ Faltas Observables/ Func. vs Lec. Cód./ P.C.E. vs. Lec. Cód. por A.S.			
Replicaciones consideradas para la pieza de conocimiento:		4 (upm01, upm03, upm04, upm05)			
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	0	1	0%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	0	1	0%	60%
Sitio		0	1	0%	0%
Protocolo experimental	Análisis	0	1	0%	
	Diseño	0	6	0%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	0%
Operacionalizaciones de los constructos	Tratamiento A	1	5	20%	
	Tratamiento B	1	5	20%	
	Efecto	0	2	0%	13%
Poblaciones	Sujetos	0	3	0%	
	Objetos experimentales	0	4	0%	0%

Tabla 13.8: Tabla de variaciones en situación experimental de replicaciones upm01, upm04, upm05

13.3.2 Generar Pieza de Conocimiento para Profesional

A partir de las tablas con información de los análisis realizados en ambos grupos de resultados (sección 13.2.3) y con las tablas de variaciones (sección 13.3.1) se empaquetan las piezas de conocimiento para el profesional. Para construir las piezas de conocimiento se usa el formulario PCR-11 (apéndice D.11). En las tablas 13.10 y 13.11 se muestra la estructura principal del formulario PCR-11 con la información de las dos piezas de conocimiento resultantes para el profesional.

De manera general la información que ofrece esta primera pieza de conocimiento (tabla 13.10) al profesional es la siguiente. De nueve replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al seleccionar el primer grupo de resultados (*upm01..05*) y apartar *upm02* la heterogeneidad es mínima por lo que se seleccionan las replicaciones *upm01* y *upm03* a *upm05* para obtener el grado de verificación logrado. Los resultados de la comparación entre ambas pruebas indican que la prueba funcional por PCE es significativamente más efectiva que la prueba de lectura de código por abstracciones sucesivas respecto a las faltas observables. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser real debido a que no se ha variado el protocolo experimental. Respecto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. La pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas en esta pieza, por último el poder de confirmación de los

Información de la partición					
Id Partición:	11.c.2				
Partición:	Efectividad/ Faltas Observables/ Func. vs Lec. Cód./ P.C.E. vs. Lec. Cód. por A.S.				
Replicaciones consideradas para la pieza de conocimiento:	3 (upm06, upm07, upm09)				
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	1	1	100%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	0	1	0%	
	Analista	1	1	100%	80%
Sitio		0	1	0%	0%
Protocolo experimental	Análisis	0	1	0%	
	Diseño	3	6	50%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	10%
Operacionalizaciones de los constructos	Tratamiento A	0	5	0%	
	Tratamiento B	0	5	0%	
	Efecto	0	2	0%	0%
Poblaciones	Sujetos	0	3	0%	
	Objetos experimentales	0	4	0%	0%

Tabla 13.9: Tabla de variaciones en situación experimental de replicaciones upm06 a upm09

resultados es del 15%.

Respecto a la segunda pieza de conocimiento (tabla 13.11), la información que ésta ofrece al profesional es la siguiente. De nueve replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al seleccionar el segundo grupo de resultados (*upm06..09*) y apartar *upm08* la heterogeneidad es mínima por lo que se seleccionan las replicaciones *upm06*, *upm07* y *upm09* para obtener el grado de verificación logrado. Los resultados de la comparación entre ambas pruebas indican que la prueba funcional por PCE es significativamente menos efectiva que la prueba de lectura de código por abstracciones sucesivas respecto a las faltas observables. El grado de verificación logrado es el siguiente, este resultado es semi-objetivo con cierta dependencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser completamente real debido a que no se ha variado la mayor parte del protocolo experimental. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas para esta pieza y el poder de confirmación de los resultados es del 18%.

Información de la Partición		
Id Partición:	11.c.1	
Partición:	Efectividad/ Faltas Observables/ Func. vs Lec. Cód./ P.C.E. vs. Lec. Cód. por A.S.	
Replicaciones de la partición:	9 (upm01..upm09)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	Inconsistencias observadas en gráfico de control y meta-análisis	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	4 (upm01, upm03, upm04, upm05)	
Comparación:	Efectividad P.C.E. > Lec. Cód. por A.S.	
Tipo de efecto global:	Significativo	
Tamaño del efecto global:	1.487 [1.2229; 1.7511]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	60%	Semi-objetivo
Sitio:	0%	Local
Protocolo experimental:	0%	Artificial
Operacionalizaciones de los constructos:	13%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	15%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 13.10: Pieza de conocimiento para profesional del primer grupo de resultados

13.3.3 Generar Pieza de Conocimiento para Investigador

De manera similar a la actividad anterior, en esta última actividad se emplean las tablas con información de los análisis realizados (sección 13.2.3) y las tablas de variaciones (sección 13.3.1). A partir de estas tablas se empaqueta las piezas de conocimiento para el investigador. Para construir estas dos piezas de conocimiento se utiliza el formulario PCR-12 (apéndice D.12). En las tablas 13.12 y 13.13 se muestra la estructura principal del formulario PCR-12 con la información de las piezas de conocimiento resultantes para el investigador.

De acuerdo a la información de esta primera pieza de conocimiento (tabla 13.12), el investigador obtiene los siguientes resultados. De la partición seleccionada se observa una heterogeneidad en el meta-análisis debido a los dos grupos de resultados observados. Al seleccionar el primer grupo (replicaciones *upm01..upm05*) se observa una heterogeneidad de resultados ocasionada por *upm02*. Al apartar esta replicación los resultados son consistentes entre sí y es posible conocer el grado de verificación alcanzado. De acuerdo a la comparación de tratamientos, la prueba funcional por PCE es significativamente más efectiva que la prueba de lectura de código por abstracciones sucesivas. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser real debido a que no se ha variado el protocolo experimental. Respecto a las

Información de la Partición		
Id Partición:	11.c.2	
Partición:	Efectividad/ Faltas Observables/ Func. vs Lec. Cód./ P.C.E. vs. Lec. Cód. por A.S.	
Replicaciones de la partición:	9 (upm01..upm09)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	Inconsistencias observadas en gráfico de control y meta-análisis	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	3 (upm06, upm07, upm09)	
Comparación:	Efectividad P.C.E. < Lec. Cód. por A.S.	
Tipo de efecto global:	Significativo	
Tamaño del efecto global:	-0.4854 [-0.7193; -0.2514]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	80%	Semi-objetivo
Sitio:	0%	Local
Protocolo experimental:	10%	Parcialmente Artificial
Operacionalizaciones de los constructos:	0%	Limitado a la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	18%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 13.11: Pieza de conocimiento para profesional del segundo grupo de resultados

operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. Los tipos de replicaciones identificados son: *Nativa*, *Interna* y *Operacional*. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 15%.

De acuerdo al análisis de causas realizado se cree que el rol del monitor así como las instancias de las faltas y el tipo de falta influyen en los resultados. Estas condiciones pueden ser testadas en replicaciones futuras.

Respecto a la información de la segunda pieza de conocimiento (tabla 13.13), el investigador obtiene los siguientes resultados. De la partición seleccionada se observa una heterogeneidad en el meta-análisis debido a los dos grupos de resultados observados. Al seleccionar el segundo grupo (replicaciones *upm06..upm09*) se observa una heterogeneidad de resultados ocasionada por *upm08*. Al apartar esta replicación los resultados son consistentes entre sí y permiten determinar el grado de verificación alcanzado. De acuerdo a la comparación de tratamientos, la prueba funcional por PCE es significativamente menos efectiva que la prueba de lectura de código por abstracciones sucesivas. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta dependencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser completamente real debido a que no se ha variado la mayor parte del protocolo experimental. Los tipos de replicaciones identificados son: *Nativa*, *Interna* y *Artificial*. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 18%.

Información de la Partición			
Id Partición:	11.c.2		
Partición:	Efectividad/ Faltas Observables/ Func. vs Lec. Cód./ P.C.E. vs. Lec. Cód. por A.S.		
Replicaciones de la partición:	9 (upm01..upm10)		
Resultados del análisis de similitudes:	Con Cambios		
Tipo de resultados encontrados:	Inconsistentes		
Información de los resultados:	Q=166.6, p<0.0001, I²=95.2%, T²= 1.119 P.C.E.: 65.36% [48.22; 82.5] f. observables: 82.48, 81.01, 82.35, 82.54, 82.57 47.1, 42.36, 40.63, 47.22 Lec. Cód. por A.S.: 48.92% [25.25; 72.59] f. observables: 47.28, 35.53, 46.31, 55.56, 54.91 53.62, 52.78, 36.81, 57.48		
Tipo de inconsistencia observada:	inconsistencias en gráfico de control y meta-análisis		
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	3 (upm06, upm07, upm09)		
Comparación:	Efectividad P.C.E. < Lec. Cód. por A.S.		
Tipo de efecto global:	Significativo		
Tamaño del efecto Global:	-0.4854 [-0.7193; -0.2514]		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	80%	Semi-objetivo	Nativa
Sitio:	0%	Local	Interna
Protocolo experimental:	10%	Parcialmente Artificial	Artificial
Operacionalizaciones de los constructos:	0%	Limitado a la forma de aplicar los tratamientos y medir los resultados	-
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes	-
Poder de confirmación de resultados	18%		
Confiabilidad de la pieza de conocimiento:	Confiable		
Análisis de Causas			
Nuevas hipótesis de trabajo:	1. El rol de medición se cree puede influir en los resultados 2. El procedimiento de medición se cree influye en los resultados 3. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados		

Tabla 13.13: Pieza de conocimiento para investigador del segundo grupo de resultados

13.4 Resultados de la Evaluación

La tipología de replicaciones propuesta ha facilitado la identificación de replicaciones hechas de forma diferenciada, nativa e interna elegidas para realizar comparaciones y obtener conocimiento de ellas. Aunque se han observado inconsistencias en los resultados de las replicaciones de esta partición, a través del *comparador de replicaciones* ha sido posible conocer el grado de verificación alcanzado en las replicaciones con resultados consistentes. Este sub-proceso también ha facilitado la identificación de variables modificadoras que se cree influyen en los resultados experimentales. Estas variables pueden ser estudiadas en profundidad en replicaciones posteriores.

Nótese que el estado de la práctica para llevar a cabo comparaciones entre replications era que sólo se comparaban los resultados experimentales y se intentaba agregarlos a través de técnicas de agregación i.e. meta-análisis, en caso de obtener resultados no reproducibles, se fracasaba en la obtención de consecuencias de la comparación de los resultados. El *comparador de replications* permite al investigador comparar las situaciones experimentales de varias replications, comparar los resultados experimentales no sólo a través del meta-análisis sino también empleando otras técnicas que facilitan examinar las variaciones de los tratamientos. Éste también facilita la identificación de variables moderadoras cuando se obtienen resultados no reproducibles. Las distintas comparaciones hechas en este sub-proceso permiten conocer el grado de verificación logrado en las replications de una partición y empaquetar este conocimiento para que sea usado por el profesional o el investigador.

Capítulo 14

Evaluación con Replicaciones Diferenciadas-Conjuntas-Externas

RESUMEN: En este capítulo se evalúan las ideas del marco para la comparación de replications referentes a las comparaciones entre replications. Esta evaluación se realiza a través del sub-proceso *comparador de replications* usando una partición compuesta por una serie de replications realizadas de forma diferenciada, conjunta y externa.

En este capítulo se evalúa el *comparador de replications* con la partición 09 descrita en la sección 10.2.7. Esta partición se compone de 13 replications (*upm01..10*, *upv05*, *uds05*, *ort05*) realizadas en años consecutivos por el grupo de investigación GrISE de la FI-UPM. La mayoría de estas replications se han realizado de manera interna, y algunas otras se han realizado en otros sitios de manera coordinada. En estas replications se han variado en diferente medida elementos de la situación experimental referentes a los experimentadores, sitio, protocolo experimental, operacionalizaciones de los constructos y poblaciones. Al igual que en la partición usada en el capítulo anterior (capítulo 13), en esta partición se compara la efectividad medida como el porcentaje de faltas observables pero para las pruebas por PCE y cobertura de decisiones.

Al contar con la tabla maestra generada en 11.1, en este capítulo se omite esta actividad. A continuación se realizan los tres sub-procesos restantes del *comparador de replications* utilizando la partición 09.

14.1 Sub-proceso Identificación de Similitudes

Este sub-proceso se compone de dos actividades que son:

- Generar tablas de particiones
- Identificar similitudes

14.1.1 Generar Tablas de Particiones

Para generar las tablas de particiones se utiliza la información de la tabla maestra referente a las replicasiones *upm01* a *upm10*, *upv05*, *uds05* y *ort05* que conforman la partición 09.

En la tabla 14.1 se muestra la tabla de partición referente a los experimentadores y sitio. Las tablas completas de esta partición se encuentran en el apéndice G.

14.1.2 Identificar Similitudes

En esta actividad del sub-proceso se usan las tablas de particiones antes descritas más el procedimiento de automatización (apéndice B) para generar las representaciones de similitudes de esta partición.

Siguiendo la automatización de cálculos, primeramente se organizan las tablas de particiones de las diferentes secciones de la situación experimental en un fichero de texto CSV¹ que representa la partición 09.

El fichero CSV de la partición 09 se ingresa como parámetro al programa *Transpose* para generar ficheros de texto que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición. Los ficheros de texto resultantes se usan para generar de forma automática las representaciones de las similitudes. Este paso se realiza tras cargar el fichero `r_script01.txt` en el entorno R (R Dev. Core Team, 2011). Este fichero contiene varias funciones en R (R Dev. Core Team, 2011) para calcular las diferentes matrices de distancias así como las funciones para realizar los análisis de conglomerados jerárquico y de escalado multidimensional.

Una vez cargado en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`, se obtienen dos ficheros EPS que representan de forma gráfica las similitudes entre las replicasiones de la partición. En la figura 14.1 se muestran los gráficos resultantes que representan las similitudes de situaciones experimentales en esta partición².

14.2 Sub-proceso Análisis de Resultados Experimentales

En este sub-proceso se comparan los resultados experimentales de las replicasiones de la partición y junto con los cambios identificados en las situaciones experimentales se determina el grado de verificación logrado en una partición. El sub-proceso *análisis de resultados experimentales* se compone de las siguientes actividades:

- Generar tabla de resultados experimentales
- Realizar comparación de resultados experimentales
- Generar tabla con información del análisis realizado

14.2.1 Generar Tabla de Resultados Experimentales

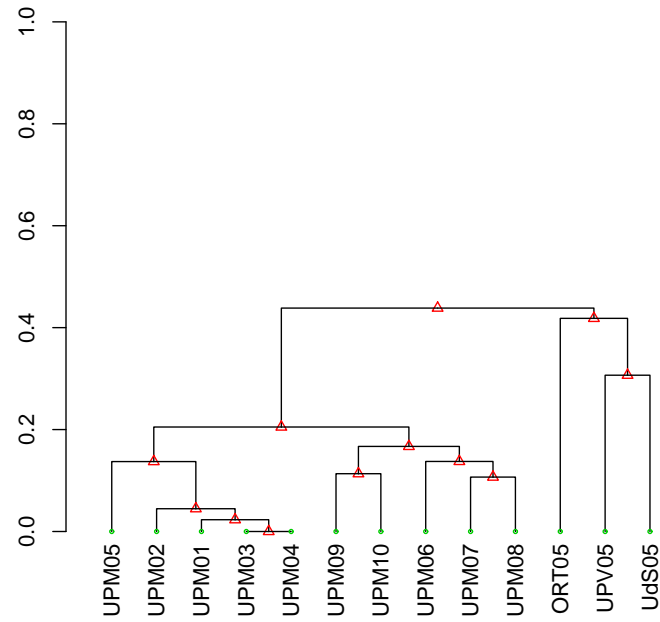
En esta actividad se extraen los resultados experimentales de las replicasiones que conforman la partición 09 y se colocan en el formulario PCR-7 (apéndice D.7). En la tabla 14.2 se muestran los resultados experimentales obtenidos de la partición 09.

¹En el primer enlace de la sección *resources* de la url <http://www.osgg.net> se encuentra disponible el fichero de texto CSV de esta partición así como los ficheros generados de forma automática que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición.

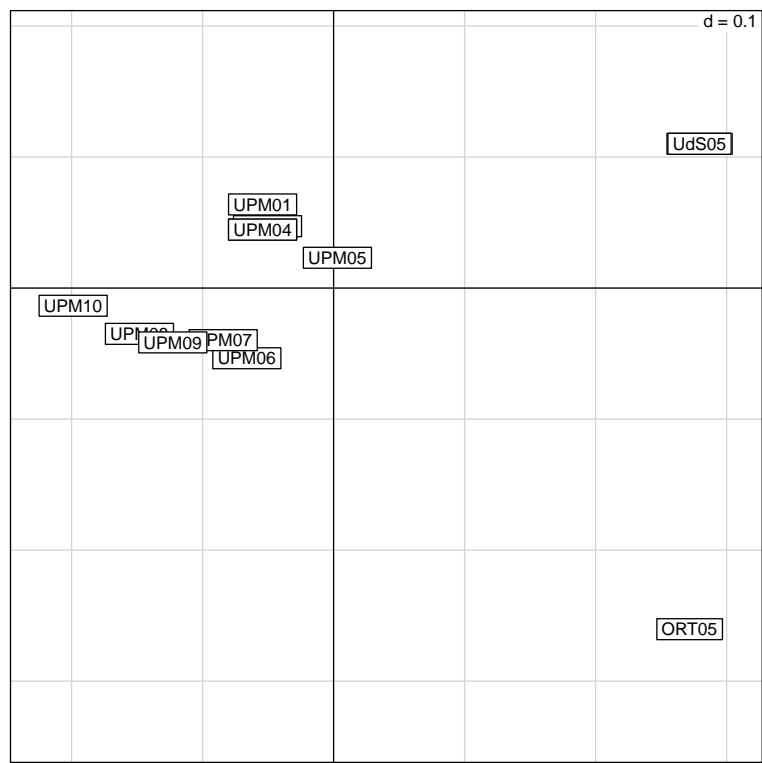
²Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script01.txt`.

Estructura	Elemento principal	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPV05	UdS05	ORT05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10
Experimentadores	Diseñador	1.1	Vegas	X	X	X	X	X	X		X	X	X	X		
		1.1	Gómez												X	X
		1.1	Solari							X		X	X	X		
	Entrenador	1.2	Juristo	X	X	X	X				X	X	X	X	X	X
		1.2	Vegas	X	X	X	X						X	X	X	X
		1.2	Letelier					X								
		1.2	Ramos						X							
		1.2	Moreno								X	X				
		1.2	Solari							X						
	Monitor	1.3	Juristo	X	X	X	X						X	X	X	X
		1.3	Vegas	X	X	X	X				X	X	X	X	X	X
		1.3	Dieste		X											
		1.3	Ferre		X											
		1.3	Moreno								X	X				
		1.3	Gómez											X	X	X
		1.3	Letelier					X								
		1.3	Ramos						X							
		1.3	Solari							X						
	Medición	1.4	Vegas	X	X	X	X									X
		1.4	Gómez											X		
		1.4	Solari					X	X	X	X	X	X		X	
	Analista	1.5	Vegas	X	X	X	X	X	X	X	X					
		1.5	Gómez											X	X	X
		1.5	Solari					X	X	X		X	X			
Sitio		2	UPM	X	X	X	X				X	X	X	X	X	X
		2	UPV					X								
		2	UdS						X							
		2	ORT							X						

Tabla 14.1: Tabla de particiones referente a los experimentadores y sitio



(a) Dendrograma



(b) MDS

Fig. 14.1: Representación de similitudes, partición 09

Id exp.	Tamaño (n_1)	Porcentaje faltas observables ($\bar{\chi}_1$)	P.C.E.		Tamaño (n_2)	Porcentaje faltas observables ($\bar{\chi}_2$)	Cob. Decisiones	
			Desv. est. (s_1)				Desv. est. (s_2)	
upm01	42	82.48	22.08		42	76.93	20.87	
upm02	39	81.01	17.75		39	85.16	19.9	
upm03	29	82.35	25.4		29	87.85	13.89	
upm04	36	82.54	16.87		36	79.83	17.6	
upm05	31	82.57	14.82		28	82.4	13.06	
upv05	31	70.05	20.13		31	69.97	24.33	
uds05	61	69.98	22.06		40	76.96	16.7	
ort05	76	76.82	23.16		76	66.35	31.98	
upm06	46	47.1	17.67		46	47.46	16.84	
upm07	48	42.36	15.36		48	47.22	21.56	
upm08	48	40.63	19.72		44	43.56	17.69	
upm09	48	47.22	15.88		50	43.67	20.73	
upm10	51	44.11	22.56		48	50	24.06	

Tabla 14.2: Resultados experimentales de la partición 09

14.2.2 Realizar Comparación de Resultados Experimentales

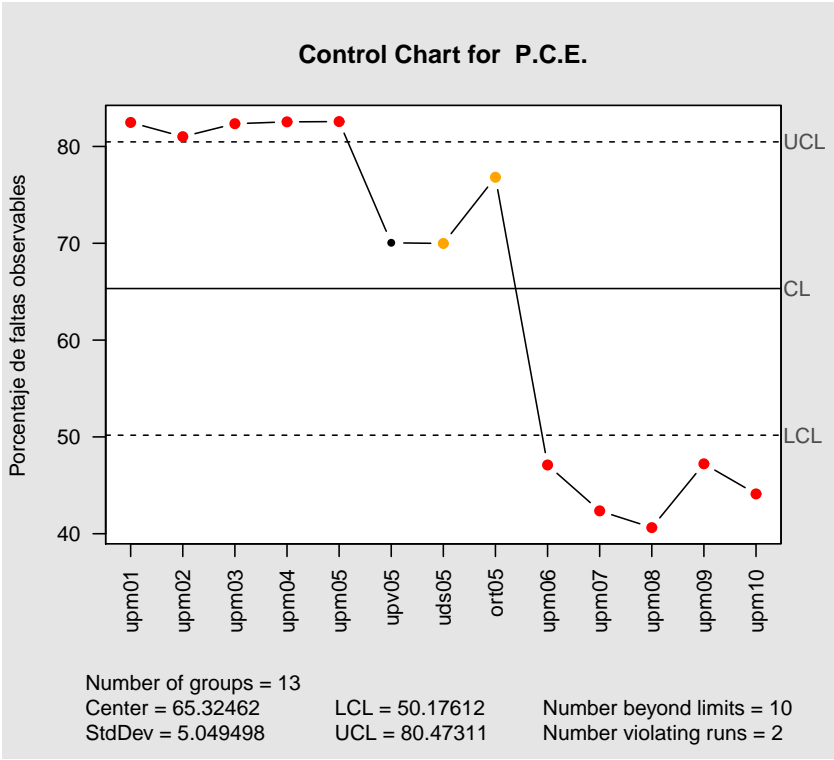
En esta actividad del sub-proceso se analizan los resultados experimentales de la partición. El análisis de los resultados se lleva a cabo a través del uso de gráficos de control y de técnicas estadísticas de agregación de resultados i.e. meta-análisis.

Para realizar estos análisis se usa la tabla de resultados experimentales de esta partición. A partir de esta tabla se generan dos gráficos de control, uno por cada tratamiento así como el meta-análisis que contiene los tamaños de efecto estimados de las replications de la partición así como el tamaño de efecto global estimado. Estos análisis se llevan a cabo cargando el fichero `r_script02.txt` en el entorno R (R Dev. Core Team, 2011) para la partición 09. Este fichero recibe como entrada la tabla de resultados experimentales de la partición. Como salida se generan tres ficheros EPS, dos ficheros que contienen los gráficos de control del par de tratamientos (PCE y cob. de decisiones), y otro fichero que contiene los resultados del meta-análisis. Con los gráficos de control y meta-análisis resultantes más las representaciones de similitudes obtenidas previamente, a continuación se examinan en conjunto estos resultados.

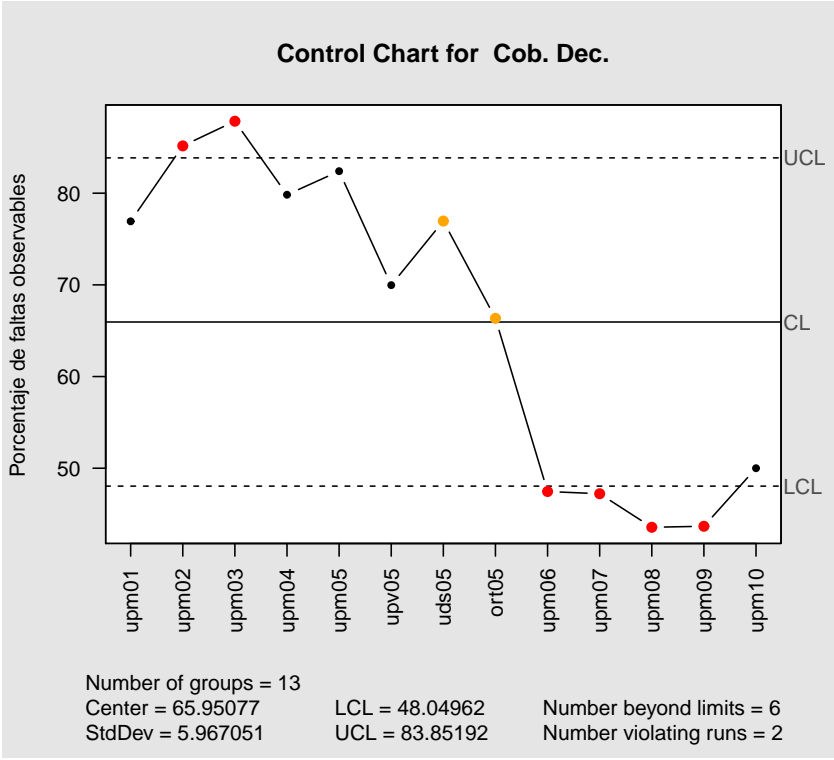
En la figura 14.2 se muestran los gráficos de control de los dos tratamientos examinados en esta partición (PCE y cob. de decisiones), mientras que en la figura 14.3 se muestran los resultados del meta-análisis³.

Respecto al gráfico de control que representa la prueba por PCE (figura 14.2a) los resultados experimentales se encuentran fuera de los límites de control. En este gráfico parecen observarse tres grupos de resultados, el primer grupo se compone por las replications *upm01* a *upm05*, el segundo grupo por las replications *upv05*, *uds05* y *ort05*, mientras que el tercer grupo por las replications *upm06* a *upm10*. A diferencia de los resultados del tercer grupo, los resultados de los dos primeros grupos se asemejan entre sí. En el gráfico de control de la prueba por cobertura de decisiones (figura 13.2b) también los resultados se encuentran fuera de los límites de control. En este gráfico se observan dos grupos de resultados, el primero formado por las replications *upm01* a *upm05*, *upv05*, *uds05* y *ort05* mientras que el segundo grupo se conforma por las replications *upm06* a *upm10*. En los resultados del meta-análisis (figura 14.3) se observa cierta heterogeneidad en la distribución de los tamaños de efecto estimados. Aunque los resultados de la prueba

³Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script02.txt`.



(a) Prueba por PCE



(b) Prueba por cob. de decisiones

Fig. 14.2: Gráficos de control resultantes

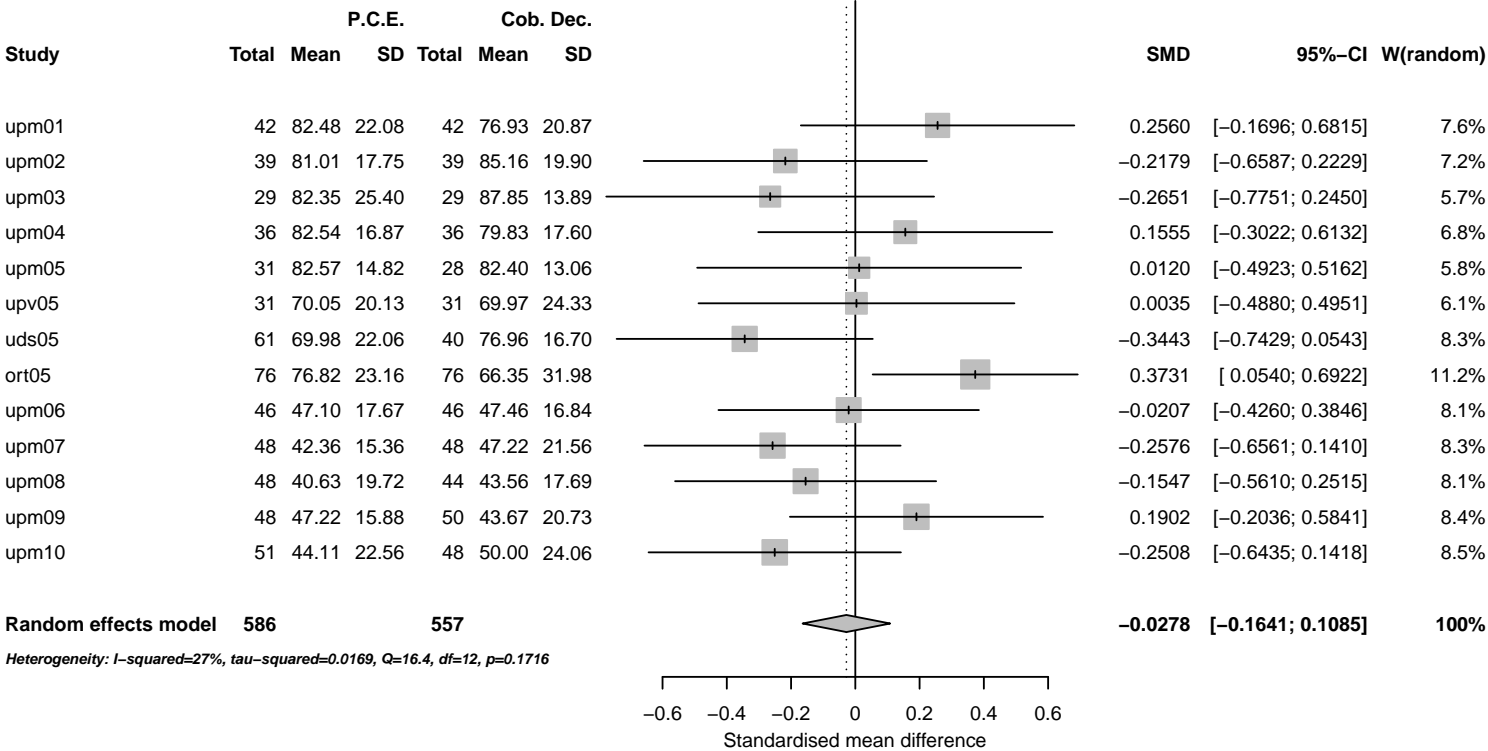


Fig. 14.3: Meta-análisis de las pruebas por PCE y cob. de decisiones

de heterogeneidad no son significativos ($Q = 16.4$, $p = 0.1716$), el índice I^2 indica un porcentaje de variabilidad total en el conjunto de tamaños de efecto estimados del 27%.

Respecto a las figuras 14.1a y 14.1b que representan las similitudes entre las replicasiones de esta partición, se observan cuatro grupos de replicasiones similares, el primer grupo se compone de las replicasiones *upm01* a *upm05*, el segundo grupo por las replicasiones *upm06* a *upm10*, el tercer grupo por *upv05* y *ort05*, mientras que el último grupo por *ort05*. Como se observa en el dendrograma, las replicasiones *ort05*, *upv05* y *uds05* son menos similares a las replicasiones *upm*, i.e. estas replicasiones tienen más cambios en los distintos elementos de la situación experimental.

De acuerdo a los resultados presentados en los gráficos de control, meta-análisis e identificación de similitudes se tienen resultados inconsistentes producto de los gráficos de control que representan la prueba por PCE y la prueba por cobertura de decisiones. Por otra parte, en los resultados de las representaciones de similitudes se observan cambios en la situación experimental. Si se utilizan todas las replicasiones de esta partición para obtener conocimiento, éste no sería confiable debido a las inconsistencias observadas. En esta situación, el investigador decide si intenta o no obtener algún conocimiento a partir del agrupamiento de resultados consistentes de la partición.

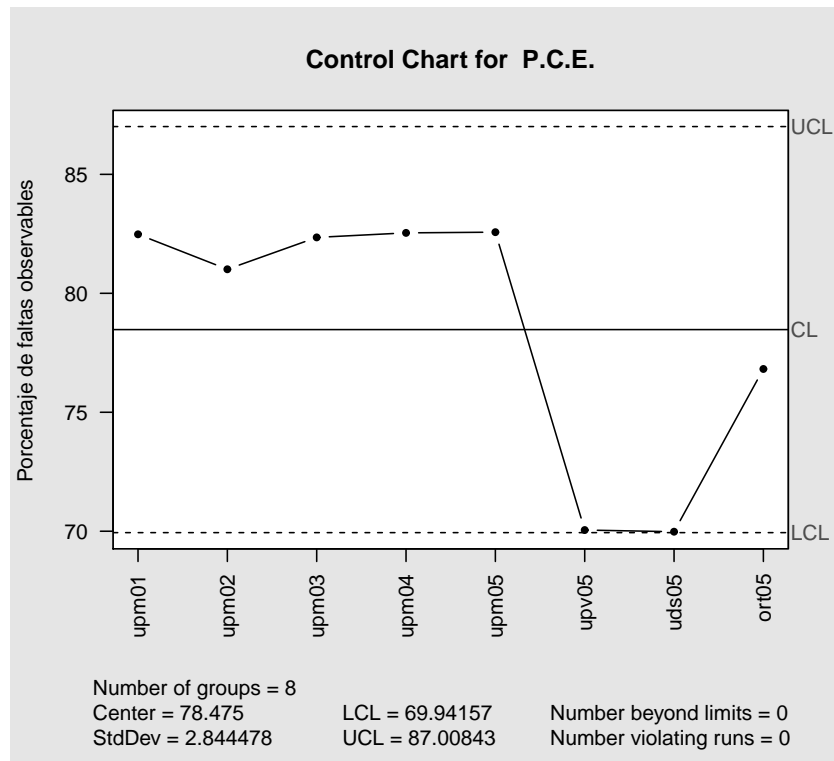
Al igual que en las particiones examinadas anteriormente, en esta partición se decide avanzar e identificar grupos de resultados consistentes con el fin obtener conocimiento. Continuando con esta actividad se examina la distribución de los tamaños de efecto tanto en los gráficos de control como en el meta-análisis. En la distribución de resultados de ambos gráficos de control se observan dos grupos de resultados. El primer grupo se compone de las replicasiones *upm01* a *upm05*, *upv05* y *uds05* mientras que el segundo grupo se compone de las replicasiones *upm06* a *upm10*. Con cada uno de estos grupos se vuelven a generar los gráficos de control y el meta-análisis para obtener resultados consistentes que permitan conocer el grado de verificación alcanzado en estos grupos. A continuación se examina cada uno de los grupos de resultados identificados.

14.2.2.1 Análisis del Primer Grupo de Resultados

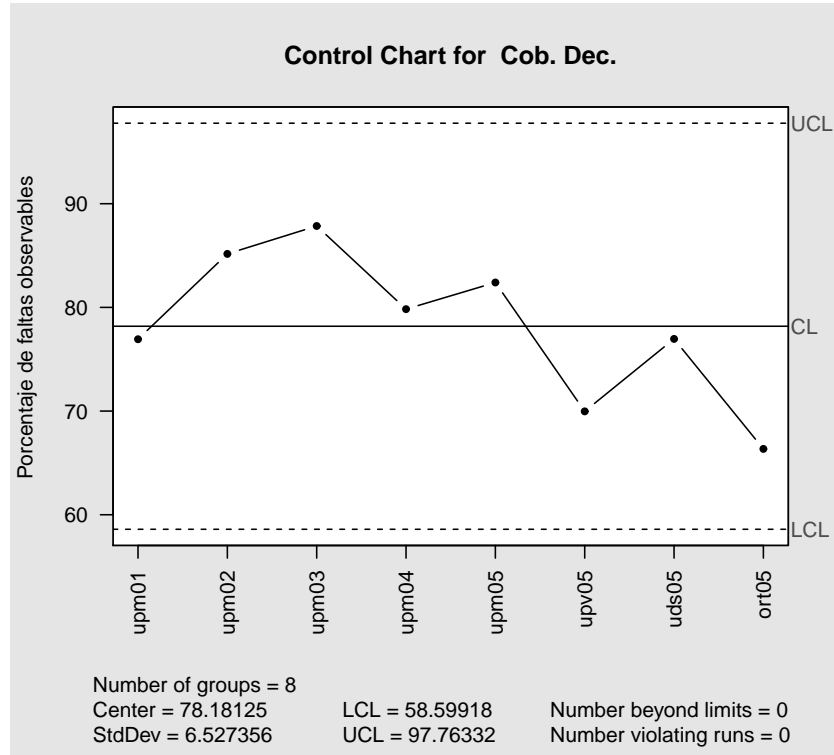
En las figuras 14.4 y 14.5 se muestran los gráficos de control y meta-análisis resultante del primer grupo de resultados.

En este primer grupo de resultados, los gráficos de control (figuras 14.4a y 14.4b) se encuentran dentro de los límites de control, sin embargo en el gráfico que representa la prueba por PCE (figura 14.4a) se observan dos resultados que están muy cercanos a los límites de control inferior. Por otra parte, en el meta-análisis se observa cierta heterogeneidad en los resultados debido en apariencia a *ort05* donde no se solapan sus intervalos como en el resto de replicasiones. Aunque la prueba de heterogeneidad no es significativa ($Q = 1 - 7$, $p = 0.1105$) el índice I^2 indica un porcentaje de variabilidad del 40.2%. En este grupo se aparta *ort05* y se usa el resto de resultados de este grupo para obtener resultados consistentes y conocer su grado de verificación logrado. Con estas replicasiones (*upm01* a *upm05*, *upv05* y *uds05*) se vuelven a generar los gráficos de control y meta-análisis. En las figuras 14.6 y 14.7 se muestran los resultados de los gráficos de control y meta-análisis tras apartar *ort05*.

Al apartar *ort05* y realizar de nuevo los gráficos de control y meta-análisis, los resultados de las replicasiones *upv05* y *uds05* se encuentran fuera de los límites de control en el gráfico que representa la prueba por P.C.E (figura 14.6a). En el gráfico que representa la prueba por cobertura de decisiones (figura 14.6b) los resultados se encuentran dentro



(a) Prueba por PCE



(b) Prueba por cob. de decisiones

Fig. 14.4: Gráficos de control del primer grupo de resultados

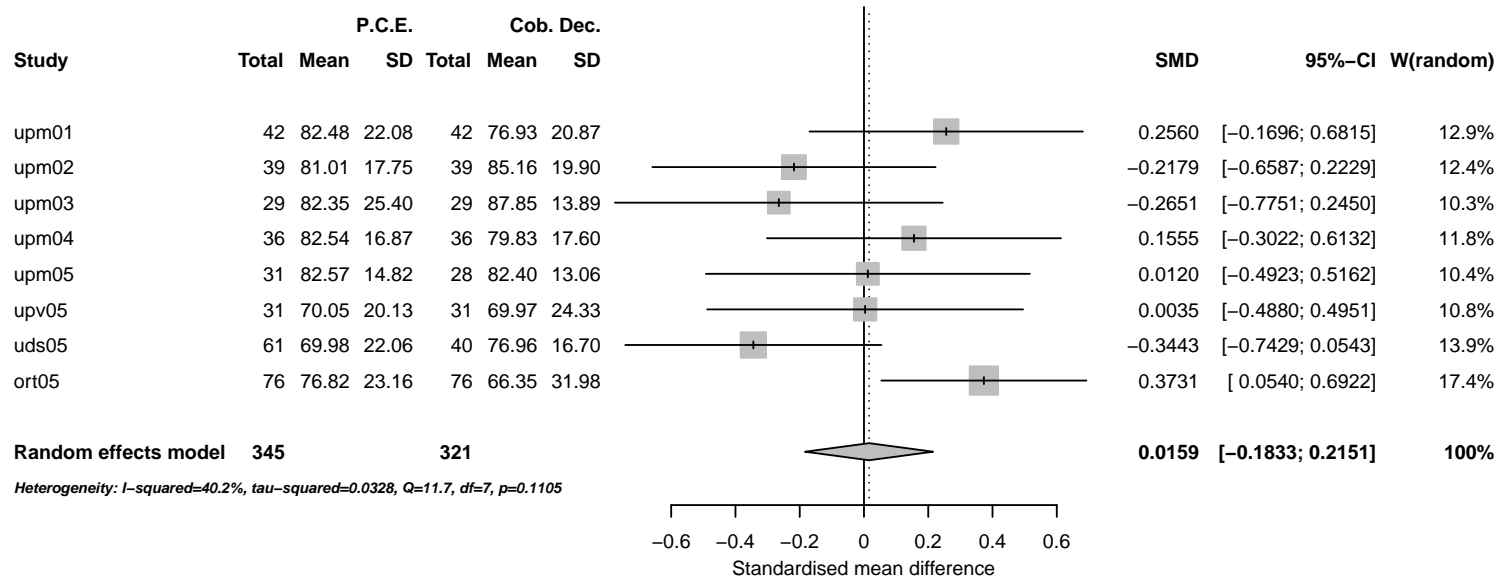
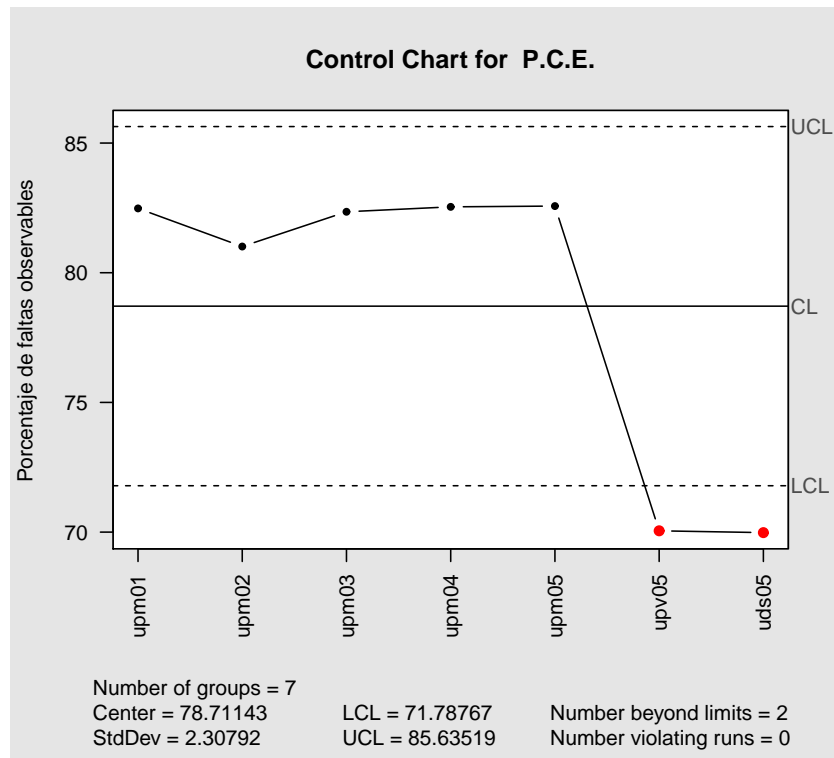
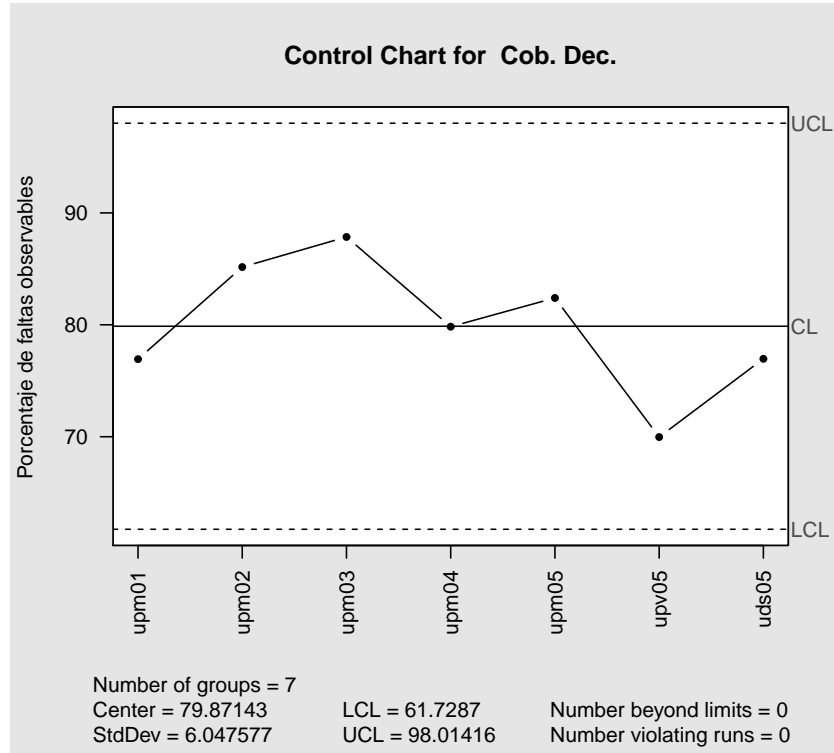


Fig. 14.5: Meta-análisis de las pruebas por PCE y cob. de decisiones del primer grupo de resultados



(a) Prueba por PCE



(b) Prueba por cob. de decisiones

Fig. 14.6: Gráficos de control del primer grupo de resultados sin ort05

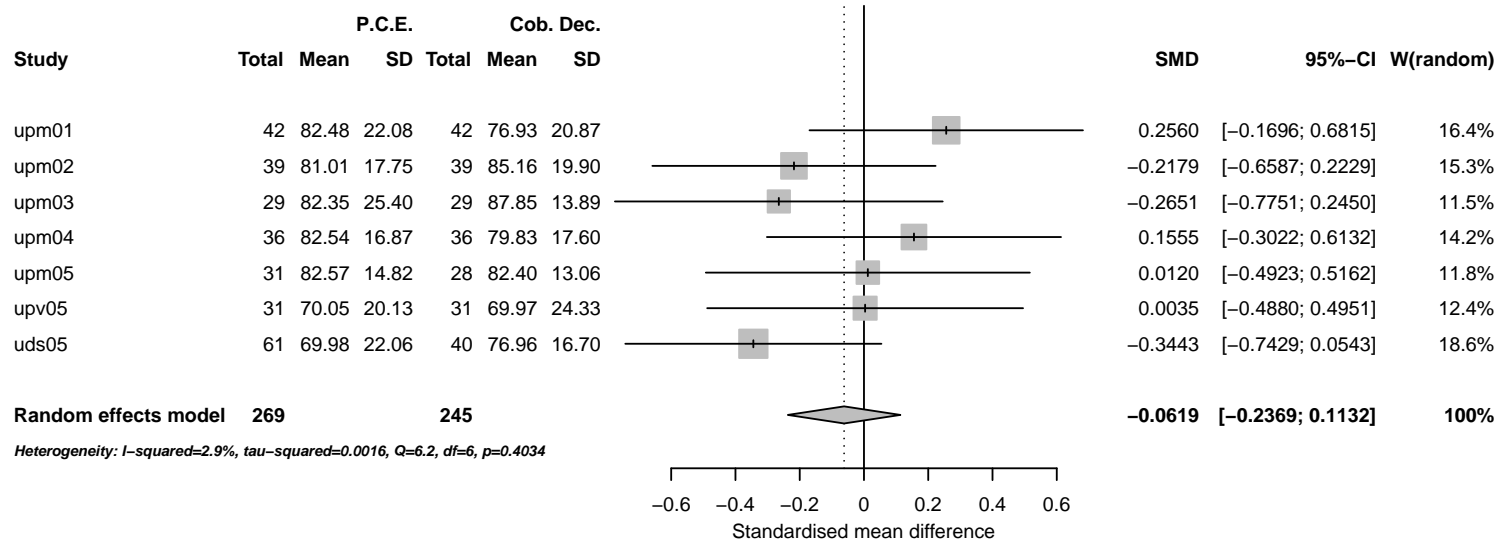


Fig. 14.7: Meta-análisis de las pruebas por PCE y cob. de decisiones del primer grupo de resultados sin ort05

de los límites de control. Respecto al meta-análisis, la prueba de heterogeneidad no es significativa ($Q = 6.2$, $p = 0.4034$), el índice I^2 indica un porcentaje de variabilidad mínimo ($I^2 = 2.9$), de manera similar, el índice T^2 indica una varianza inter-estudios mínima. Al observar resultados fuera de los límites de control en uno de los gráficos, se apartan las replicaciones *upv05* así como *uds05* y se vuelven a generar los gráficos de control y meta-análisis con el resto de replicaciones de este primer grupos de resultados. En las figuras 14.8 y 14.9 se muestran los resultados de los gráficos de control y meta-análisis tras apartar *upv05* y *uds05*.

Tras realizar de nuevo los gráficos de control y meta-análisis, los resultados en ambos gráficos se encuentran dentro de los límites de control, respecto al meta-análisis no se observa heterogeneidad en la distribución de los tamaños de efecto estimados. Los resultados de estos nuevos análisis son consistentes y pueden usarse para determinar el grado de verificación logrado así como empaquetar este conocimiento.

14.2.2.2 Análisis del Segundo Grupo de Resultados

Ya que se identificaron resultados consistentes en el primer grupo de resultados, a continuación se examina el segundo grupo. En las figuras 14.10 y 14.11 se muestran los gráficos de control y meta-análisis de este segundo grupo de resultados.

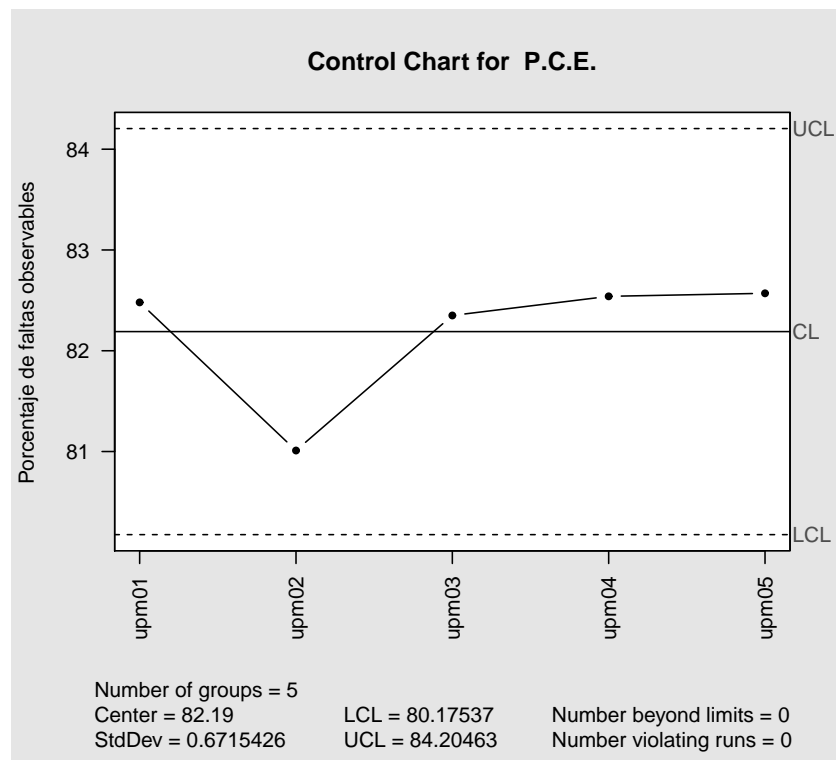
En este segundo grupo de resultados, ambos gráficos se encuentran dentro de los límites de control (figuras 14.10a y 14.10b). Por otra parte, los resultados del meta-análisis (figura 14.11) no indican heterogeneidad en la distribución de tamaños de efecto estimados, la prueba de heterogeneidad no es significativa ($Q = 3.5$, $p = 0.4825$), de manera similar los resultados de los índices I^2 y T^2 ambos son igual a cero.

Cuando se apartan replicaciones para intentar obtener conocimiento es recomendable llevar a cabo el análisis de causas para tratar de identificar los elementos de la situación experimental que se cree influyen en los resultados experimentales. A continuación se realiza el análisis de causa en cada uno de los grupos de resultados (intra-grupos) y por último a nivel de grupos de resultados (inter-grupos).

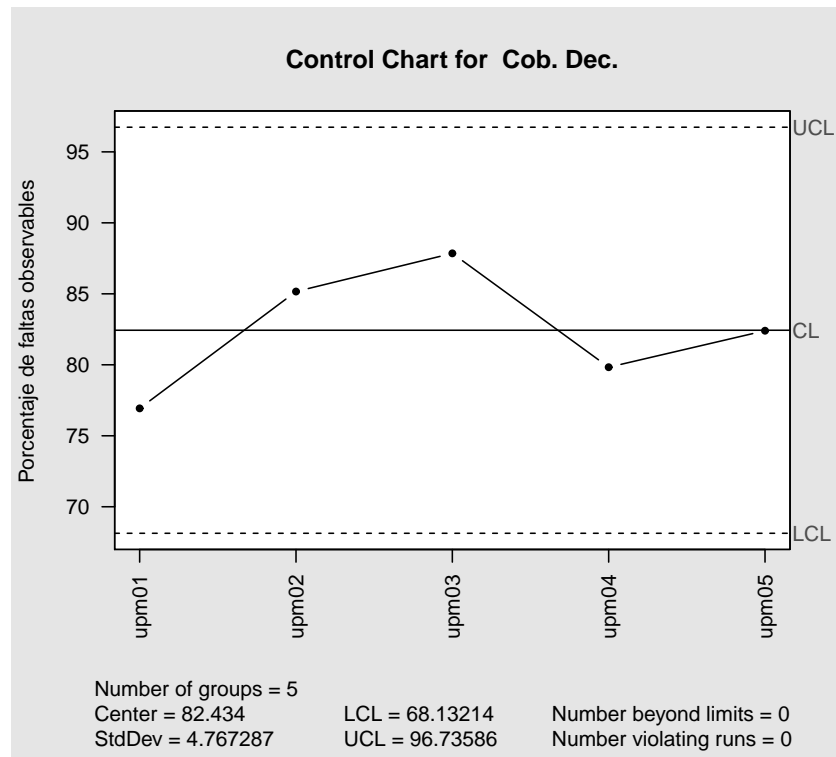
14.2.2.3 Análisis de Causas del Primer Grupo de Resultados

En las figuras 14.12, 14.13, 14.14 y 14.15 se muestran los diagramas de Ishikawa (Ishikawa, 1990) con el número de cambios identificados en los elementos de la situación experimental de este primer grupo de resultados (replicaciones *upm01* a *upm05*, *upv05*, *uds05* y *ort05*). Por otra parte en la tabla 14.3 se muestran las variaciones en los diferentes elementos de la situación experimental.

De acuerdo a los resultados observados en los gráficos de control y meta-análisis, las replicaciones *upm01* a *upm05* presentan resultados consistentes en la métrica usada en esta partición (porcentaje de faltas observables). Los cambios identificados entre las replicaciones parecen no influir de forma significativa en los resultados. Por el contrario parece que alguno de los cambios identificados en *upv05*, *uds05* y *ort05* influyó en la efectividad de las pruebas examinadas en esta partición. Al igual que en las replicaciones *upm01* a *upm05* en estas tres replicaciones se usaron las mismas instancias y tipos de faltas, sin embargo la efectividad en estas replicaciones es menor, este decremento se observa mejor en la prueba por PCE (figura 14.4a). Como se observa en la tabla 14.3 en estas replicaciones se han realizado varios cambios que dificulta identificar alguno como posible causante del decremento en la efectividad de las pruebas.



(a) Prueba por PCE



(b) Prueba por cob. de decisiones

Fig. 14.8: Gráficos de control del primer grupo de resultados sin ort05, upv05 y uds05

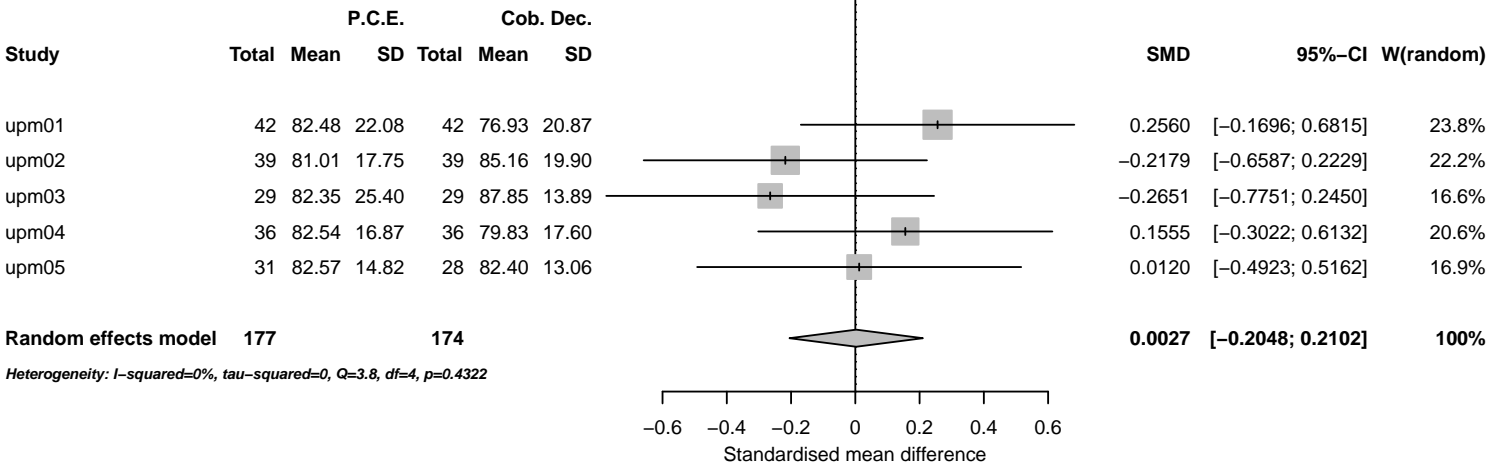
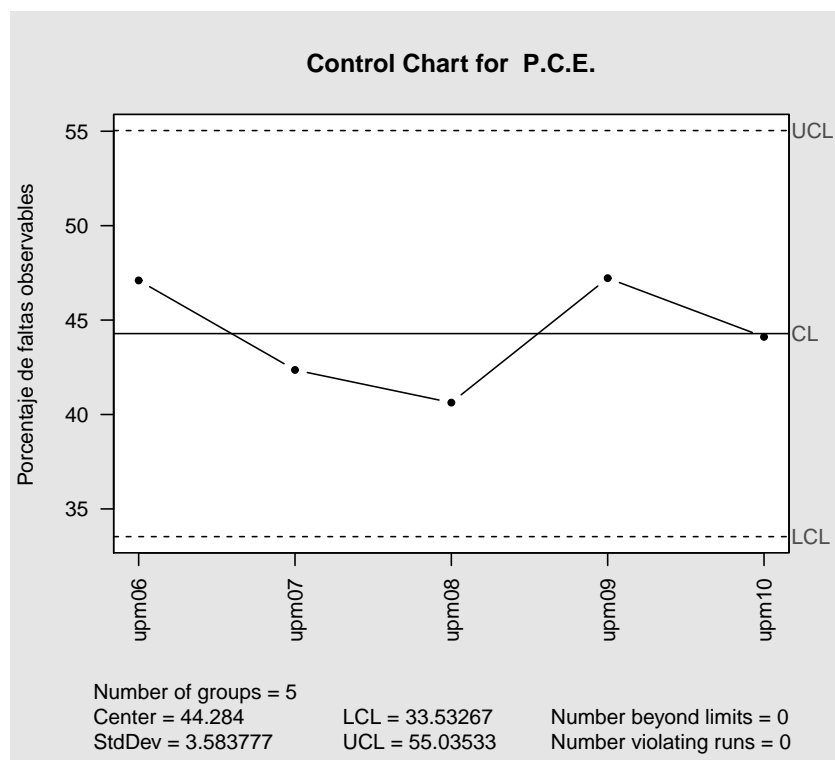
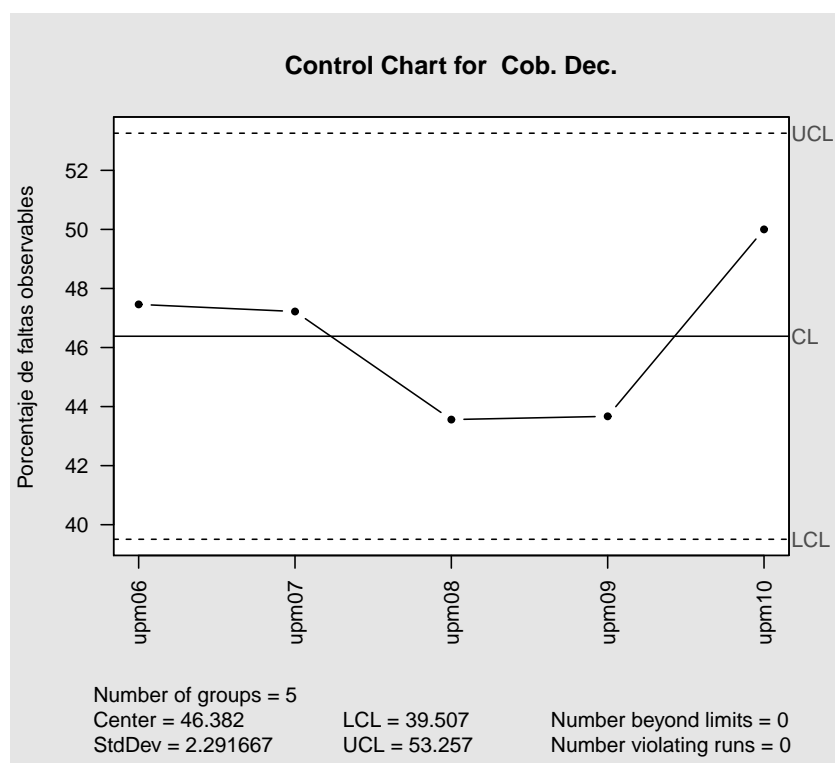


Fig. 14.9: Meta-análisis de las pruebas por PCE y cob. de decisiones del primer grupo de resultados sin ort05, upv05 y uds05



(a) Prueba por PCE



(b) Prueba por cob. de decisiones

Fig. 14.10: Gráficos de control del segundo grupo de resultados

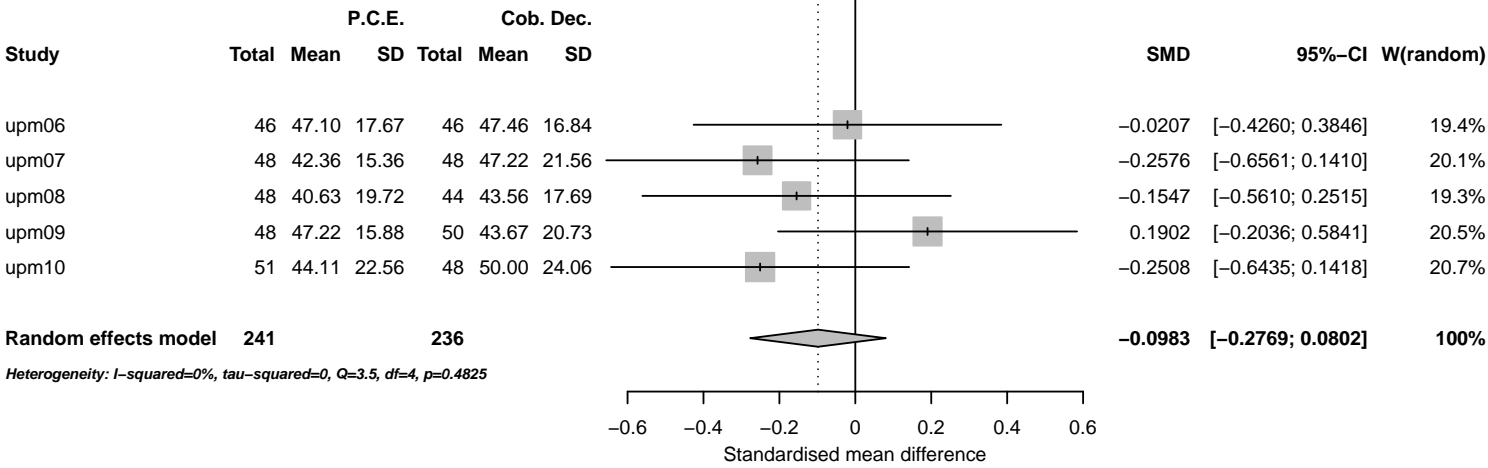


Fig. 14.11: Meta-análisis de las pruebas por PCE y cob. de decisiones del segundo grupo de resultados

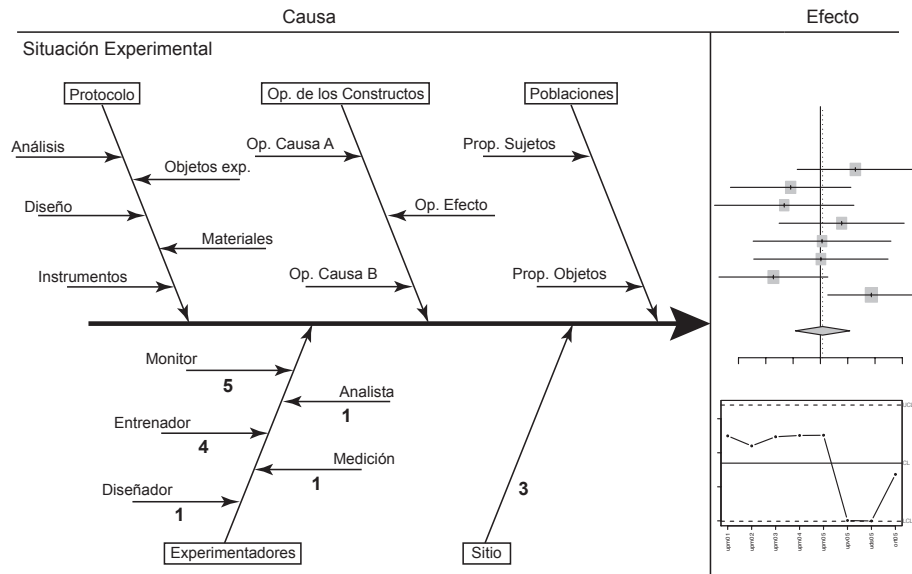


Fig. 14.12: Cambios identificados en experimentadores y sitio

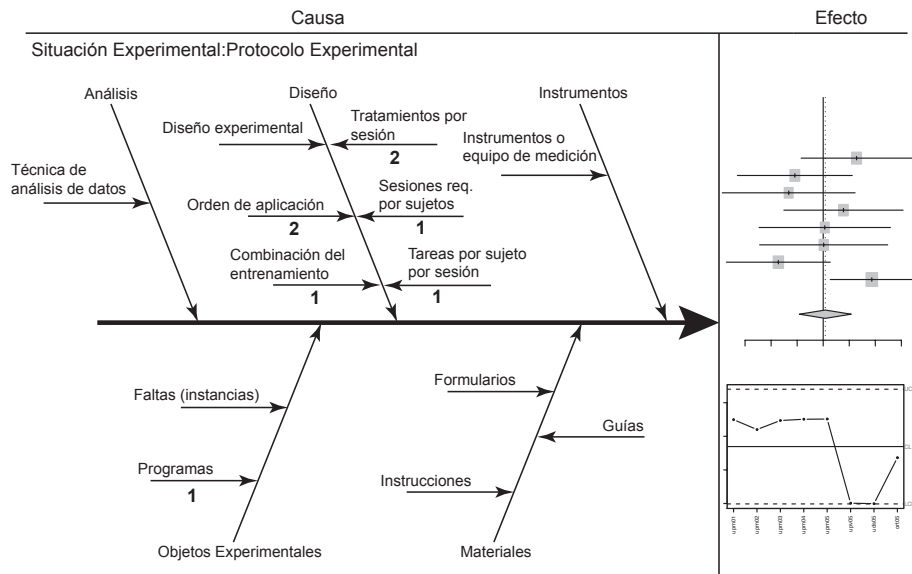


Fig. 14.13: Cambios identificados en el protocolo experimental

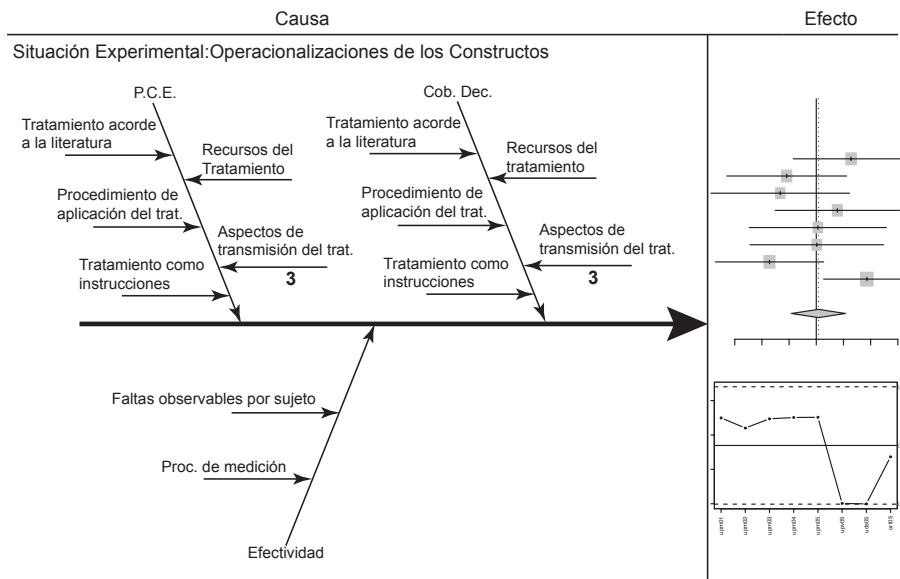


Fig. 14.14: Cambios identificados en operacionalizaciones de los constructos

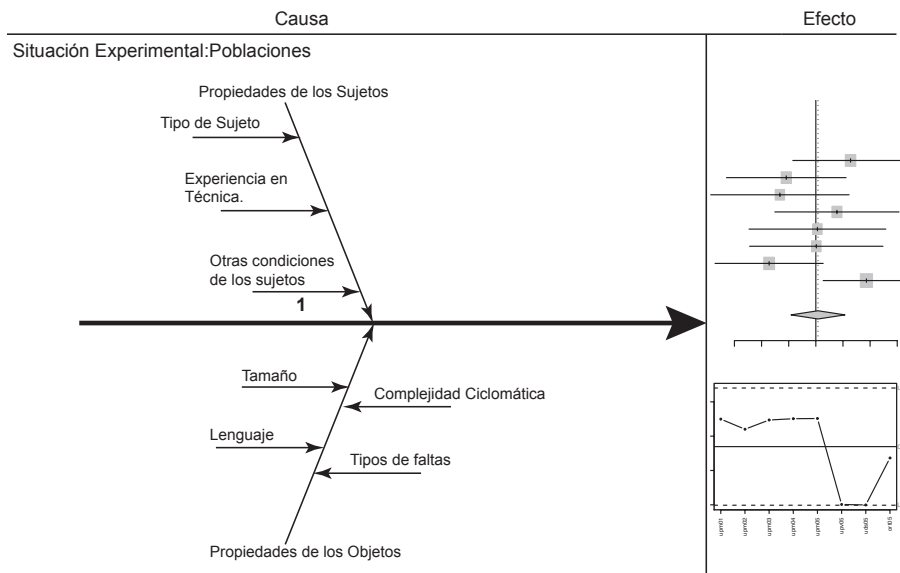


Fig. 14.15: Cambios identificados en poblaciones

Elemento principal	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPV05	UdS05	ORT05	UPM05
Diseñador	1.1	Vegas	X	X	X	X	X	X		X
	1.1	Solari							X	
Entrenador	1.2	Juristo	X	X	X	X				X
	1.2	Vegas	X	X	X	X				
	1.2	Letelier					X			
	1.2	Ramos						X		
	1.2	Moreno								X
	1.2	Solari							X	
Monitor	1.3	Juristo	X	X	X	X				
	1.3	Vegas	X	X	X	X				X
	1.3	Dieste		X						
	1.3	Ferre		X						
	1.3	Moreno								X
	1.3	Letelier					X			
	1.3	Ramos						X		
	1.3	Solari							X	
Medición	1.4	Vegas	X	X	X	X				
	1.4	Solari					X	X	X	X
Analista	1.5	Vegas	X	X	X	X	X	X	X	X
	1.5	Solari					X	X	X	
	2	UPM	X	X	X	X				X
	2	UPV					X			
	2	UdS						X		
	2	ORT							X	
Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X	X			X	X
	3.2.2	E , F					X			
	3.2.2	E, F, R						X		
Combinación del entrenamiento con las sesiones	3.2.3	Intercalado					X	X		
	3.2.3	Secuencial	X	X	X	X			X	X
Tratamientos por sesión	3.2.4	3 Técnicas, 1 Programa	X	X	X	X				X
	3.2.4	1 Técnica, 3 Programas					X	X		
	3.2.4	2 Técnicas, 2 Progrmas							X	
Sesiones requeridas por sujeto para aplicar tratamientos	3.2.5	1 Sesión	X	X	X	X			X	X
	3.2.5	2 Sesiones					X	X		
Tareas por sujetopor sesión	3.2.6	1 Técnica, 1 Programa	X	X	X	X	X	X		X
	3.2.6	2 Técnicas, 2 Programas							X	
Programas	3.4.1	Cmdline	X	X	X	X	X	X	X	X
	3.4.1	Nametbl	X	X	X	X	X	X	X	X
	3.4.1	Ntree	X	X	X	X	X	X		X
Aspectos de transmisión del tratamiento (P.C.E.)	4.1.1.5	Con experiencia en curso universitario	X	X	X	X			X	X
	4.1.1.5	Sin experiencia en curso universitario	X							
	4.1.1.5	Con experiencia en tutorial					X			
	4.1.1.5	Sin experiencia en tutorial						X		
Aspectos de transmisión del tratamiento (Cob. Dec.)	4.1.2.5	Con experiencia en curso universitario	X	X	X	X			X	X
	4.1.2.5	Sin experiencia en curso universitario	X							
	4.1.2.5	Con experiencia en tutorial					X			
	4.1.2.5	Sin experiencia en tutorial						X		
Otras condiciones de los sujetos	5.1.3	Conocimientos de prog aptos p/ exp.	X	X	X	X	X	X		X
	5.1.3	Conocimiento limitado de programación							X	

Tabla 14.3: Cambios identificados en los elementos de la situación experimental del primer grupo

Respecto a la replicación *ort05*, se emplearon sujetos con un conocimiento limitado en programación (alumnos comenzando estudios de pregrado). Respecto a los gráficos de control, en esta replicación se observa mayor efectividad en la prueba por PCE que en la prueba por cob. de decisiones. Este decremento de efectividad en la prueba por cob. de decisiones tal vez se deba a que se requiere mayor conocimiento de programación para llevar a cabo esta prueba. En el caso de *upv05* y *uds05* respecto a los distintos cambios identificados se cree que la combinación del entrenamiento con las sesiones así como las sesiones requeridas por sujeto para aplicar los tratamientos pudieron influir en los resultados. En estas dos replicaciones no se disponía de tiempo suficiente para realizar, en primer lugar, el entrenamiento durante el curso académico y posteriormente la replicación. Ante esta situación se intercaló el entrenamiento con las sesiones de la replicación, después de entrenar a los sujetos en alguna de las pruebas se realizó la sesión del experimento replicado referente a esa prueba. Debido al tiempo que disponían los sujetos para aplicar los tratamientos, los sujetos llevaron a cabo la aplicación de cada tratamiento en dos sesiones. Se cree que el usar dos sesiones por tratamiento pudo influir en los resultados.

De estos cambios identificados se definen las siguientes hipótesis que pueden ser testadas en replicaciones futuras:

- Utilizar sujetos con conocimiento limitado de programación se cree influye en los resultados de la prueba por cob. de decisiones.
- Combinar el entrenamiento con las sesiones del experimento se cree influye en los resultados del experimento.
- Usar más de una sesión por sujeto para aplicar el tratamiento debido a restricciones de tiempo se cree influye en los resultados del experimento.

14.2.2.4 Análisis de Causas Inter-grupos de Resultados

El segundo grupo de resultados no requiere de análisis de causas porque los resultados en este grupo son consistentes, los cambios identificados en las replicaciones *upm06* a *upm10* parecen no influir en los resultados. A continuación se analizan los cambios entre estos dos grupos de resultados, en las figuras 14.16 y 14.17 se muestran los diagramas de Ishikawa (Ishikawa, 1990) con el número de cambios identificados entre estos dos grupos de resultados, por otra parte en la tabla 14.4 se muestran los elementos variados entre estos dos grupos.

Los cambios identificados entre estos dos grupos de resultados se relacionan con las instancias de faltas de los objetos experimentales y con los tipos de faltas. En las replicaciones *upm01* a *upm05*, *upv05*, *uds05* y *ort05* se utilizaron dos versiones de cada programa. Cada versión tiene cierto número de faltas inyectadas en distintos lugares del código fuente. De manera similar en las replicaciones *upm06* a *upm10* se usó una sola versión de los programas con faltas inyectadas en lugares distintos a las versiones usadas en las replicaciones *upm01* a *upm05*, *upv05*, *uds05* y *ort05*. Se cree que las faltas empleadas en ambos grupos de replicaciones influyen en los resultados. Este cambio puede ser testado en futuras replicaciones para verificar esta hipótesis.

Una vez completada esta actividad del sub-proceso *análisis de resultados experimentales* se seleccionan los dos grupos de resultados consistentes para intentar empaquetar las piezas de conocimiento. A continuación se desarrolla la última actividad de este sub-proceso.

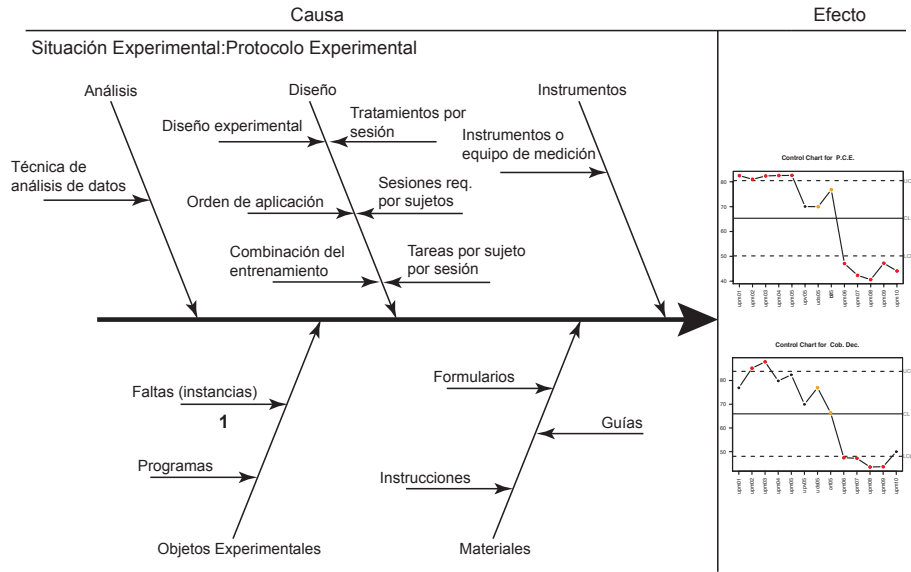


Fig. 14.16: Cambios identificados en el protocolo experimental

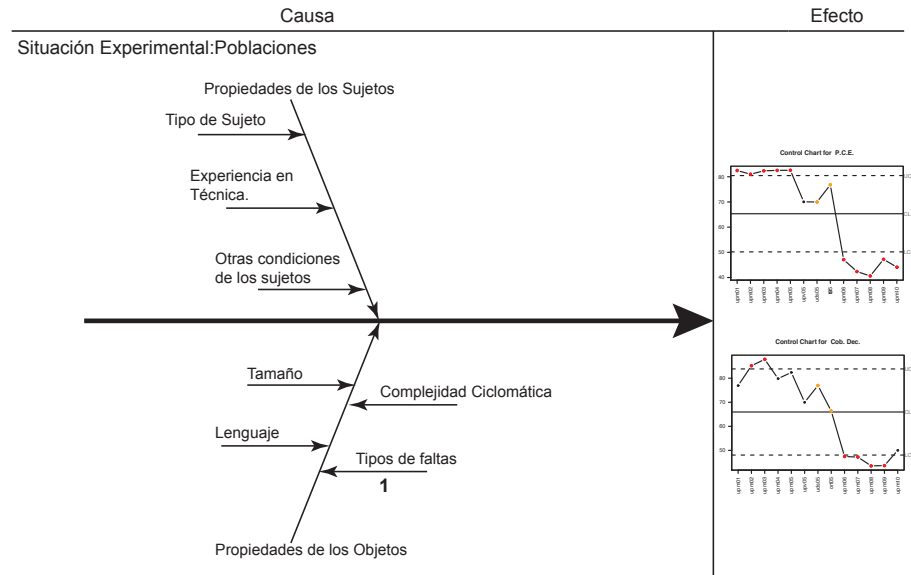


Fig. 14.17: Cambios identificados en poblaciones

Elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPV05	UdS05	ORT05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10
Faltas (instancias)	3.4.2	Instancia G	X	X	X	X	X	X	X	X					
	3.4.2	Instancia H	X	X	X	X	X	X	X	X					
	3.4.2	Instancia I									X	X	X	X	X
Tipos de Faltas	5.2.4	Grupo de Faltas G	X	X	X	X	X	X	X	X					
	5.2.4	Grupo de Faltas H	X	X	X	X	X	X	X	X					
	5.2.4	Grupo de Faltas I									X	X	X	X	X

Tabla 14.4: Cambios identificados en los elementos de la situación experimental a nivel de inter-grupos

14.2.3 Generar Tabla con Información del Análisis Realizado

Una vez realizada la comparación de resultados experimentales, en esta actividad se coloca en el formulario PCR-9 (apéndice D.9) la información de las comparaciones de resultados realizadas en la actividad anterior. En este caso se usan dos formularios, uno por cada grupo de resultados consistentes. En las tablas 14.5 y 14.6 se muestra la estructura principal del formulario PCR-9 con la información de los resultados de las comparaciones realizadas en la actividad anterior. Los grupos de resultados consistentes se han identificado como particiones 09.1 y 09.2.

14.3 Sub-proceso Empaquetado de Piezas de Conocimiento

Como se ha descrito en los capítulos anteriores, ese sub-proceso se compone de las actividades:

- Generar tabla de variaciones en situación experimental
- Generar pieza de conocimiento para profesional
- Generar pieza de conocimiento para investigador

14.3.1 Generar Tabla de Variaciones en Situación Experimental

En esta actividad del sub-proceso se usa la información de las tablas de particiones generadas en el sub-proceso *identificación de similitudes* para calcular los cambios realizados en la situación experimental de las replicaciones que constituyen los resultados consistentes, en este caso las replicaciones del primer grupo de resultados *upm01* a *upm05*,

Información de la Partición	
Id. Partición:	09.1
Partición:	Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.
Replicaciones de la partición:	13 (upm01..upm10, upv05, uds05, ort05)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes
Información de los resultados:	Q=16.4, p=0.1716, I ² =27%, T ² = 0.0169 P.C.E.: 65.32% [50.18; 80.47] f. observables: 82.48, 81.01, 82.35, 82.54, 82.57 70.05, 69.98, 76.82, 47.1, 42.36, 40.63, 47.22, 44.11 Cob. Dec.: 65.95% [48.05; 83.85] f. observables: 76.93, 85.16, 87.85, 79.83, 82.4 69.97, 76.96, 66.35, 47.46, 47.22, 43.56, 43.67, 50
Tipo de inconsistencia observada:	Inconsistencias en gráficos de control
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	5 (upm01..upm05)
Comparación del par de tratamientos:	P.C.E. = Cob. Dec. (Sin diferencias en efectividad)
Tipo de efecto global:	No significativo
Tamaño del efecto global:	0.0027 [-0.2048; 0.2102]

Tabla 14.5: Información de las comparaciones de resultados del primer grupo de resultados

Información de la Partición	
Id. Partición:	09.2
Partición:	Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.
Replicaciones de la partición:	13 (upm01..upm10, upv05, uds05, ort05)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes
Información de los resultados:	Q=16.4, p=0.1716, I ² =27%, T ² = 0.0169 P.C.E.: 65.32% [50.18; 80.47] f. observables: 82.48, 81.01, 82.35, 82.54, 82.57 70.05, 69.98, 76.82, 47.1, 42.36, 40.63, 47.22, 44.11 Cob. Dec.: 65.95% [48.05; 83.85] f. observables: 76.93, 85.16, 87.85, 79.83, 82.4 69.97, 76.96, 66.35, 47.46, 47.22, 43.56, 43.67, 50
Tipo de inconsistencia observada:	Inconsistencias en gráficos de control
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	5 (upm06..upm10)
Comparación del par de tratamientos:	P.C.E. = Cob. Dec. (Sin diferencias en efectividad)
Tipo de efecto global:	No significativo
Tamaño del efecto global:	-0.0983 [-0.2769; 0.0802]

Tabla 14.6: Información de las comparaciones de resultados del segundo grupo de resultados

upv05, *uds05* y *ort05* así como las replicaciones del segundo grupo de resultados *upm06* a *upm10*. En esta actividad se utiliza el formulario PCR-10 (apéndice D.10) para colocar los cálculos de los cambios realizados. La tabla resultante se utiliza en las siguientes actividades del sub-proceso para empaquetar las piezas de conocimiento. En las tablas 14.7 y 14.8 se muestra la estructura principal del formulario PCR-10 con la el cálculo de variaciones en ambos grupos de resultados consistentes.

Información de la partición					
Id Partición:		09.1			
Partición:		Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.			
Replicaciones consideradas para la pieza de conocimiento:		5 (upm01..upm05)			
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	0	1	0%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	0	1	0%	60%
Sitio		0	1	0%	0%
Protocolo experimental	Análisis	0	1	0%	
	Diseño	0	6	0%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	0%
Operacionalizaciones de los constructos	Tratamiento A	1	5	20%	
	Tratamiento B	1	5	20%	
	Efecto	0	2	0%	13%
Poblaciones	Sujetos	0	3	0%	
	Objetos experimentales	0	4	0%	0%

Tabla 14.7: Tabla de variaciones en situación experimental de replicaciones upm01 a upm05

14.3.2 Generar Pieza de Conocimiento para Profesional

A partir de las tablas con información de los análisis realizados en ambos grupos de resultados (sección 14.2.3) y con las tablas de variaciones (sección 14.3.1) se determina el grado de verificación logrado en esa partición y se empaqueta este conocimiento para que sea usado por el profesional. Para empaquetar este conocimiento se usa el formulario PCR-11 (apéndice D.11). En las tablas 14.9 y 14.10 se muestra la estructura principal del formulario PCR-11 con la información de las dos piezas de conocimiento resultantes para el profesional.

De manera general la información que ofrece esta primera pieza de conocimiento (tabla 14.9) al profesional es la siguiente. De 13 replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al seleccionar el primer grupo de resultados (*upm01..05*, *upv05*, *uds05* y *ort05*) y apartar las replicaciones *upv05*, *uds05* y *ort05* la heterogeneidad es mínima por lo que se seleccionan las replicaciones *upm01* a *upm05* para obtener el grado de verificación logrado. Los resultados de la comparación entre la prueba por PCE y la prueba por cobertura de decisiones indican que no hay diferencia significativa en la efectividad de ambas pruebas respecto a las faltas observables. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser real debido a que no se ha variado el protocolo experimental. Respecto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. La pieza de conocimiento es confiable al tener

Información de la partición					
Id Partición:	09.2				
Partición:	Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.				
Replicaciones consideradas para la pieza de conocimiento:	5 (upm06..upm10)				
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	1	1	100%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	1	1	100%	100%
Sitio		0	1	0%	0%
Protocolo experimental	Análisis	0	1	0%	
	Diseño	3	6	50%	
	Instrumentos	0	1	0%	
	Objetos experimentales	0	2	0%	
	Materiales	0	3	0%	10%
Operacionalizaciones de los constructos	Tratamiento A	0	5	0%	
	Tratamiento B	0	5	0%	
	Efecto	1	2	50%	17%
Poblaciones	Sujetos	1	3	33%	
	Objetos experimentales	0	4	0%	17%

Tabla 14.8: Tabla de variaciones en situación experimental de replicaciones upm06 a upm10

una baja heterogeneidad en las replicaciones consideradas en esta pieza, por último el poder de confirmación de los resultados es del 15%.

Respecto a la segunda pieza de conocimiento (tabla 14.10), la información que ésta ofrece al profesional es la siguiente. De 13 replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al seleccionar el segundo grupo de resultados (*upm06..10*) la heterogeneidad es mínima por lo que se seleccionan todas las replicaciones de este segundo grupo para obtener el grado de verificación logrado. Los resultados de la comparación entre la prueba por PCE y la prueba por cobertura de decisiones indican que no hay diferencia significativa en la efectividad de ambas pruebas respecto a las faltas observables. El grado de verificación logrado es el siguiente, este resultado es objetivo independiente de los experimentadores y local dependiente del sitio. El resultado observado puede no ser completamente real debido a que no se ha variado la mayor parte del protocolo experimental. En cuanto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. Respecto a las poblaciones, se tiene cierto conocimiento sobre algunas propiedades de las poblaciones usadas en estas replicaciones. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas para esta pieza y el poder de confirmación de los resultados es del 29%.

Información de la Partición		
Id Partición:	09.1	
Partición:	Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.	
Replicaciones de la partición:	13 (upm01..upm10, upv05, uds05, ort05)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	Inconsistencias observadas en gráficos de control	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	5 (upm01..upm05)	
Comparación:	Efectividad en P.C.E. = Cob. Dec. (Sin diferencias en efectividad)	
Tipo de efecto global:	No significativo	
Tamaño del efecto global:	-0.0983 [-0.2769; 0.0802]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	60%	Semi-objetivo
Sitio:	0%	Local
Protocolo experimental:	0%	Artificial
Operacionalizaciones de los constructos:	13%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	15%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 14.9: Pieza de conocimiento para profesional del primer grupo de resultados

14.3.3 Generar Pieza de Conocimiento para Investigador

De manera similar a la actividad anterior, en esta última actividad se emplean las tablas con información de los análisis realizados (sección 14.2.3) y las tablas de variaciones (sección 14.3.1). A partir de estas tablas se empaquetan las piezas de conocimiento para el investigador. Para construir estas dos piezas de conocimiento se utiliza el formulario PCR-12 (apéndice D.12). En las tablas 14.11 y 14.12 se muestra la estructura principal del formulario PCR-12 con la información de las piezas de conocimiento resultantes para el investigador.

De acuerdo a la información de esta primera pieza de conocimiento (tabla 14.11), el investigador obtiene los siguientes resultados. De la partición seleccionada se observa una heterogeneidad en el meta-análisis debido a los dos grupos de resultados observados. Al seleccionar el primer grupo (replicaciones *upm01..upm05*, *upv05*, *uds05* y *ort05*) se observa una heterogeneidad de resultados ocasionada por *upv05*, *uds05* y *ort05*. Al apartar estas replicaciones los resultados son consistentes entre sí y se puede conocer el grado de verificación logrado. De acuerdo a la comparación de tratamientos la prueba por PCE y la prueba por cobertura de decisiones indican que no hay diferencia significativa en la efectividad respecto a las faltas observables. El grado de verificación logrado es el siguiente, este resultado es parcialmente objetivo con cierta independencia de los experimentadores y local dependiente del sitio. El resultado observado puede no ser real debido a que no se ha variado el protocolo experimental. Respecto a las operacionalizaciones, se

Información de la Partición		
Id Partición:	09.2	
Partición:	Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.	
Replicaciones de la partición:	13 (upm01..upm10, upv05, uds05, ort05)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	Inconsistencias observadas en gráficos de control	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	5 (upm06..upm10)	
Comparación:	Efectividad en P.C.E. = Cob. Dec. (Sin diferencias en efectividad)	
Tipo de efecto global:	No significativo	
Tamaño del efecto global:	-0.0983 [-0.2769; 0.0802]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	100%	Objetivo
Sitio:	0%	Local
Protocolo experimental:	10%	Parcialmente Artificial
Operacionalizaciones de los constructos:	17%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	17%	Con cierto conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	29%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 14.10: Pieza de conocimiento para profesional del segundo grupo de resultados

tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. Los tipos de replicaciones identificados son: *Nativa*, *Interna* y *Operacional*. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 15%.

De acuerdo al análisis de causas realizado las siguientes condiciones se cree influyen en los resultados experimentales:

- Usar sujetos con conocimiento limitado de programación
- Combinar el entrenamiento con las sesiones del experimento
- Usar más de una sesión por sujeto para aplicar el tratamiento debido a restricciones de tiempo
- Las instancias de las faltas usadas así como el tipo de falta

Respecto a la información de la segunda pieza de conocimiento (tabla 14.12), el investigador obtiene los siguientes resultados. De la partición seleccionada se observa una heterogeneidad en el meta-análisis debido a los dos grupos de resultados observados. Al seleccionar el segundo grupo (replicaciones *upm06..upm10*) los resultados son consistentes entre sí y es posible conocer el grado de verificación alcanzado. De acuerdo a la comparación de tratamientos la prueba por PCE y la prueba por cobertura de decisiones indican que no hay diferencia significativa en la efectividad respecto a las faltas observables. Este resultado es objetivo independiente de los experimentadores y local dependiente del sitio. El grado de verificación logrado es el siguiente, el resultado observado puede no ser com-

Información de la Partición			
Id Partición:	09.1		
Partición:	Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.		
Replicaciones de la partición:	13 (upm01..upm10, upv05, uds05, ort05)		
Resultados del análisis de similitudes:	Con Cambios		
Tipo de resultados encontrados:	Inconsistentes		
Información de los resultados:	Q=16.4, p=0.1716, I²=27%, T²= 0.0169 P.C.E.: 65.32% [50.18; 80.47] f. observables: 82.48, 81.01, 82.35, 82.54, 82.57 70.05, 69.98, 76.82, 47.1, 42.36, 40.63, 47.22, 44.11 Cob. Dec.: 65.95% [48.05; 83.85] f. observables: 76.93, 85.16, 87.85, 79.83, 82.4 69.97, 76.96, 66.35, 47.46, 47.22, 43.56, 43.67, 50		
Tipo de inconsistencia observada:	Inconsistencias en gráficos de control		
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	5 (upm01..upm05)		
Comparación:	Efectividad en P.C.E. = Cob. Dec. (Sin diferencias en efectividad)		
Tipo de efecto global:	No Significativo		
Tamaño del efecto Global:	-0.0983 [-0.2769; 0.0802]		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	60%	Semi-objetivo	Nativa
Sitio:	0%	Local	Interna
Protocolo experimental:	0%	Artificial	-
Operacionalizaciones de los constructos:	13%	Con cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados	Operacional
Poblaciones:	0%	Sin conocimiento de poblaciones equivalentes	-
Poder de confirmación de resultados	15%		
Confiabilidad de la pieza de conocimiento:	Confiable		
Análisis de Causas			
Nuevas hipótesis de trabajo:	1. Utilizar sujetos con conocimiento limitado de programación se cree influye en los resultados de la prueba por Cob. de Decisiones 2. Combinar el entrenamiento con las sesiones del experimento se cree influye en los resultados del experimento 3. Usar más de una sesión por sujeto para aplicar el tratamiento debido a restricciones de tiempo se cree influye en los resultados del experimento 4. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados		

Tabla 14.11: Pieza de conocimiento para investigador del primer grupo de resultados

pletamente real debido a que no se ha variado la mayor parte del protocolo experimental. En cuanto a las operacionalizaciones, se tiene cierto conocimiento en la forma de aplicar los tratamientos y medir los resultados. Respecto a las poblaciones, se tiene cierto conocimiento sobre algunas propiedades de las poblaciones usadas en estas replicaciones. Los tipos de replicaciones identificados son: *Nativa*, *Interna*, *Artificial*, *Operacional* y *Poblacional*. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 29%.

Información de la Partición	
Id Partición:	09.2
Partición:	Efectividad/ Faltas Observables/ Func. vs Estruct./ P.C.E. vs. Cob. Dec.
Replicaciones de la partición:	13 (upm01..upm10, upv05, uds05, ort05)
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes
Información de los resultados:	Q=16.4, p=0.1716, I ² =27%, T ² = 0.0169 P.C.E.: 65.32% [50.18; 80.47] f. observables: 82.48, 81.01, 82.35, 82.54, 82.57 70.05, 69.98, 76.82, 47.1, 42.36, 40.63, 47.22, 44.11 Cob. Dec.: 65.95% [48.05; 83.85] f. observables: 76.93, 85.16, 87.85, 79.83, 82.4 69.97, 76.96, 66.35, 47.46, 47.22, 43.56, 43.67, 50
Tipo de inconsistencia observada:	inconsistencias en gráficos de control

Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	5 (upm06..upm10)		
Comparación:	Efectividad en P.C.E. = Cob. Dec. (Sin diferencias en efectividad)		
Tipo de efecto global:	No significativo		
Tamaño del efecto Global:	-0.0983 [-0.2769; 0.0802]		

Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	100%	Objetivo	Nativa
Sitio:	0%	Local	Interna
Protocolo experimental:	10%	Parcialmente Artificial	Artificial
Operacionalizaciones de los constructos:	17%	Limitado a la forma de aplicar los tratamientos y medir los resultados	Operacional
Poblaciones:	17%	Sin conocimiento de poblaciones equivalentes	Poblacional
Poder de confirmación de resultados	29%		
Confiabilidad de la pieza de conocimiento:	Confiable		
Análisis de Causas			
Nuevas hipótesis de trabajo:	1. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados		

Tabla 14.12: Pieza de conocimiento para investigador del segundo grupo de resultados

De acuerdo al análisis de causas realizado las instancias de las faltas así como el tipo de falta se cree influyen en los resultados. Estas condiciones pueden ser testadas en replicaciones posteriores.

14.4 Resultados de la Evaluación

La tipología de replicaciones propuesta ha facilitado la identificación de replicaciones hechas de forma diferenciada, conjunta y externa elegidas para realizar comparaciones y obtener conocimiento de ellas. Aunque se han observado inconsistencias en los resultados de las replicaciones de esta partición, a través del *comparador de replicaciones* ha sido posible conocer el grado de verificación alcanzado en las replicaciones con resultados consistentes. Este sub-proceso también ha facilitado la identificación de variables mode-

radoras que se cree influyen en los resultados experimentales. Estas variables pueden ser estudiadas en profundidad en replicaciones posteriores.

Nótese que el estado de la práctica para llevar a cabo comparaciones entre replicaciones era que sólo se comparaban los resultados experimentales y se intentaba agregarlos a través de técnicas de meta-análisis, en caso de obtener resultados no reproducibles, se fracasaba en la obtención de consecuencias de la comparación de los resultados. El *comparador de replicaciones* permite al investigador comparar las situaciones experimentales de varias replicaciones, comparar los resultados experimentales no sólo a través del meta-análisis sino también empleando otras técnicas que facilitan examinar las variaciones de los tratamientos. Éste también facilita la identificación de variables moderadoras cuando se obtienen resultados no reproducibles. Las distintas comparaciones hechas en este subproceso permiten conocer el grado de verificación logrado en las replicaciones de una partición y empaquetar este conocimiento para que sea usado por el profesional o el investigador.

Capítulo 15

Evaluación con Replicaciones Diferenciadas-Ajenas-Externas

RESUMEN: En este capítulo se evalúan las ideas del marco para la comparación de replications referentes a las comparaciones entre replications. Esta evaluación se realiza a través del sub-proceso *comparador de replications* usando una partición compuesta por una serie de replications de tipo diferenciada, ajena y externa.

En este capítulo se emplean las particiones 01, 02 y 03 (sección 10.2.7) que conforman la partición 48. Esta partición tiene en común los constructos de causa que son las pruebas Funcionales y Estructurales. La finalidad de evaluar esta partición es conocer si los resultados se mantienen al usar particiones que comparten un mismo constructo de causa pero con variaciones en las operacionalizaciones de causa. Esta partición se compone de las replications *upm01* a *upm11* y *upv05* (partición 01), *umd82* a *umd84* (partición 02) y las replications *ukl94*, *ukl95* así como *uos97* (partición 03). En esta partición se comparara la efectividad medida como el porcentaje de faltas observadas de las pruebas funcional y estructural.

El *comparador de replications* se compone de la actividad *generar tabla maestra* y de los sub-procesos: *identificación de similitudes*, *análisis de resultados experimentales* y *empaquetado de piezas de conocimiento*.

Al contar con la tabla maestra generada en 11.1, en este capítulo se omite esta actividad. A continuación se realizan los tres sub-procesos restantes del *comparador de replications* usando la partición 48.

15.1 Sub-proceso Identificación de Similitudes

Este sub-proceso se compone de dos actividades que son:

- Generar tablas de particiones
- Identificar similitudes

15.1.1 Generar Tablas de Particiones

Para generar las tablas de particiones se usa la información de la tabla maestra referente a las replicasiones *upm01* a *upm11* y *upv05* (partición 01), *umd82* a *umd84* (partición 02) y las replicasiones *ukl94*, *ukl95* así como *uos97* (partición 03) que conforman esta partición.

En la tabla 15.1 se muestra la tabla de partición referente a los experimentadores y sitio. Las tablas completas de esta partición se encuentran en el apéndice G.

15.1.2 Identificar Similitudes

En esta actividad del sub-proceso se utilizan las tablas de particiones antes descritas más el procedimiento de automatización (apéndice B) para generar las representaciones de similitudes de esta partición.

Siguiendo la automatización de cálculos, primeramente se organizan las tablas de particiones de las diferentes secciones de la situación experimental en un fichero de texto CSV¹ que representa la partición 48.

El fichero CSV de la partición 48 se ingresa como parámetro al programa *Transpose* para generar ficheros de texto que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición. Los ficheros de texto resultantes se utilizan para generar de forma automática las representaciones de las similitudes. Este paso se realiza tras cargar el fichero *r_script01.txt* en el entorno R (R Dev. Core Team, 2011). Este fichero contiene varias funciones en R (R Dev. Core Team, 2011) para calcular las diferentes matrices de distancias así como las funciones para realizar los análisis de conglomerados jerárquico y de escalado multidimensional.

Una vez cargado en el entorno R (R Dev. Core Team, 2011) el fichero *r_script01.txt*, se obtienen dos ficheros EPS que representan de forma gráfica las similitudes entre las replicasiones de la partición. En la figura 15.1 se muestran los gráficos resultantes que representan las similitudes de situaciones experimentales en esta partición².

15.2 Sub-proceso Análisis de Resultados Experimentales

En este sub-proceso se comparan los resultados experimentales de las replicasiones de la partición y junto con los cambios identificados en las situaciones experimentales se determina el grado de verificación logrado en una partición. El sub-proceso *análisis de resultados experimentales* se compone de las siguientes actividades:

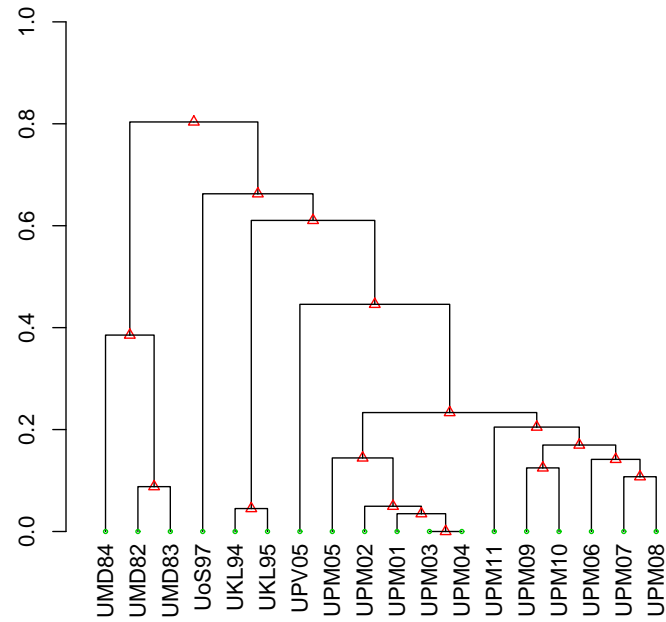
- Generar tabla de resultados experimentales
- Realizar comparación de resultados experimentales
- Generar tabla con información del análisis realizado

¹En el primer enlace de la sección *resources* de la url <http://www.osgg.net> se encuentra disponible el fichero de texto CSV de esta partición así como los ficheros generados de forma automática que representan las tablas de contingencia de los (sub)elementos de la situación experimental de esta partición.

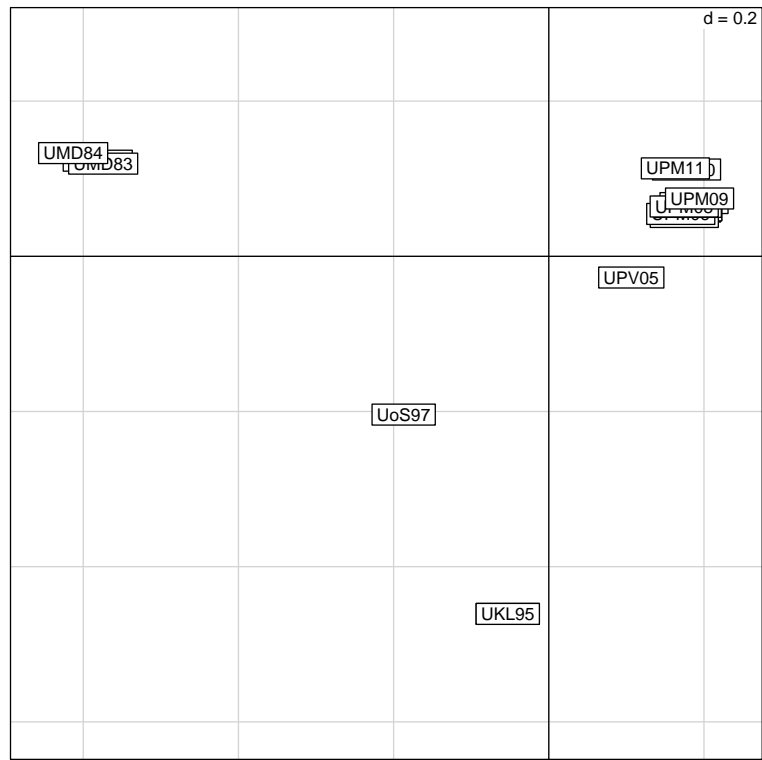
²Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero *r_script01.txt*.

Estructura	Elemento principal	ID	Categoría	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04	UPV05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Experimentadores	Diseñador	1.1	Basili	X	X	X															
		1.1	Selby	X	X	X															
		1.1	Kamsties				X	X													
		1.1	Lott				X	X													
		1.1	Roper						X												
		1.1	Wood						X												
		1.1	Miller						X												
		1.1	Vegas							X	X	X	X	X	X	X	X	X			X
		1.1	Gómez																X	X	
		1.1	Solari													X	X	X			
		1.1	Apa																		X
	Entrenador	1.2	Basili	X	X	X															
		1.2	Selby	X	X	X															
		1.2	Kamsties				X	X													
		1.2	Lott				X	X													
		1.2	Roper						X												
		1.2	Wood						X												
		1.2	Miller						X												
		1.2	Juristo							X	X	X	X		X	X	X	X	X	X	X
		1.2	Vegas							X	X	X	X				X	X	X	X	X
		1.2	Letellier											X							
		1.2	Moreno												X	X					
	Monitor	1.3	Basili	X	X	X															
		1.3	Selby	X	X	X															
		1.3	Kamsties				X	X													
		1.3	Lott				X	X													
		1.3	Roper						X												
		1.3	Wood						X												
		1.3	Miller						X												
		1.3	Juristo							X	X	X	X				X	X	X	X	
		1.3	Vegas							X	X	X	X		X	X	X	X	X	X	X
		1.3	Dieste								X										
		1.3	Ferre								X										
		1.3	Moreno												X	X					
		1.3	Gómez															X	X	X	X
		1.3	Letellier											X							
		1.3	Apa																		X

Tabla 15.1: Tabla de particiones referente a los experimentadores y sitio



(a) Dendrograma



(b) MDS

Fig. 15.1: Representación de similitudes, partición 48

15.2.1 Generar Tabla de Resultados Experimentales

En esta actividad se extraen los resultados experimentales de las replicaciones que conforman la partición 48 y se colocan en el formulario PCR-7 (apéndice D.7). En la tabla 12.2 se muestran los resultados experimentales obtenidos en esta partición³.

Id exp.	P.C.E.			Cob. Decisiones		
	Tamaño (n_1)	Porcentaje faltas observadas ($\bar{\chi}_1$)	Desv. est. (s_1)	Tamaño (n_2)	Porcentaje faltas observadas ($\bar{\chi}_2$)	Desv. est. (s_2)
umd82	29	60.68	23.18	29	44.73	22.77
umd83	13	41.89	20.33	13	34.22	21
umd84	32	47.89	14.36	32	34.82	19.29
ukl94	27	51.01	17.55	27	51.87	22.66
ukl95	21	61.52	21.06	18	55.55	21.36
uos97	47	55.2	15.36	47	57.48	20.28
upm00	48	37.26	20.3	46	35.5	17.31
upm01	42	62.4	22.88	42	57.59	19.56
upm02	39	81.45	17.7	39	84.88	19.56
upm03	29	75.78	13.39	29	70.27	19.19
upm04	36	66.6	17.72	36	63.38	18.64
upm05	32	62.2	23.08	32	63.39	23.5
upv05	14	43.36	21.25	10	54.02	21.45
upm06	46	42.75	17.09	46	35.51	20.37
upm07	48	38.89	15.5	48	40.97	23.56
upm08	46	37.68	18.07	40	28.75	19.24
upm09	48	42.71	19.72	50	38	24.06
upm10	25	29.33	16.85	25	31.33	21.14
upm11	51	47.71	19.15	48	42.01	27.926

Tabla 15.2: Resultados experimentales de la partición 48

15.2.2 Realizar Comparación de Resultados Experimentales

En esta actividad del sub-proceso se analizan los resultados experimentales de la partición. El análisis de los resultados se lleva a cabo a través del uso de gráficos de control y de técnicas estadísticas de agregación de resultados i.e. meta-análisis.

Para realizar estos análisis se usa la tabla de resultados experimentales de esta partición. A partir de esta tabla se generan dos gráficos de control, uno por cada tratamiento así como el meta-análisis que contiene los tamaños de efecto estimados de las replicaciones de la partición así como el tamaño de efecto global estimado. Estos análisis se llevan a cabo cargando el fichero `r_script02.txt` en el entorno R (R Dev. Core Team, 2011) para la partición 48. Este fichero recibe como entrada la tabla de resultados experimentales de la partición. Como salida se generan tres ficheros EPS, dos ficheros que contienen los gráficos de control del par de tratamientos (prueba funcional y prueba estructural), y otro fichero que contiene los resultados del meta-análisis. Con los gráficos de control y meta-análisis resultantes más las representaciones de similitudes obtenidas previamente, a continuación se examinan en conjunto estos resultados.

En las figuras 15.2 se muestran los gráficos de control de los dos tratamientos examinados en esta partición (prueba funcional y prueba estructural), mientras que en la figura

³Las métricas de las replicaciones *umd82* a *umd84*, y *uos97* se transformaron a las mismas métricas del resto de replicaciones para mantener una misma métrica, el *porcentaje de faltas observadas por sujeto*.

15.3 se muestran los resultados del meta-análisis⁴.

Al analizar los resultados se observa que en ambos gráficos de control la replicación *upm02* se encuentra fuera de los límites de control. Respecto a los resultados del meta-análisis, aunque la prueba de heterogeneidad no es significativa ($Q = 24.3$, $p = 0.1116$) se observa cierto porcentaje de variabilidad total en los tamaños de efecto estimados ($I^2 = 30\%$).

En el dendrograma de la figura 15.1a y en el MDS de la figura 15.1b se observan grupos de replicaciones que se unen o se encuentran a diferentes distancias. Entre mayor es la distancia en que se unen los grupos, la similitud entre éstos es menor. Por ejemplo en el dendrograma de la figura 15.1a el grupo formado por las replicaciones *umd82* a *umd84* se une a una distancia de 0.8 del resto de replicaciones i.e. hay más diferencias en los elementos de la situación experimental entre este grupo y el resto. Por otra parte, las replicaciones *upm* son más similares entre sí, se observan dos grupos que se unen a una distancia poco mayor a 0.2 i.e. en estas replicaciones son pocas las variaciones que hay en los elementos de la situación experimental.

De acuerdo a los resultados presentados en los gráficos de control, meta-análisis e identificación de similitudes se tienen resultados inconsistentes producto principalmente de los gráficos de control. En los resultados de la identificación de similitudes se observan cambios en la situación experimental. Si se utilizan todas las replicaciones de esta partición para obtener conocimiento, éste no sería confiable debido a las inconsistencias observadas. En esta situación, el investigador decide si intenta o no generar piezas de conocimiento a partir del agrupamiento de resultados consistentes de la partición.

Con el fin de obtener algún conocimiento de esta partición se decide avanzar e identificar grupos de resultados consistentes. Siguiendo con esta actividad, a continuación se examina la distribución de los tamaños de efecto tanto en los gráficos de control como en el meta-análisis. En la distribución de resultados de los gráficos de control que representan las pruebas funcional y estructural se observa que *upm02* se encuentra fuera de los límites de control, esta replicación se aparta del resto y se vuelven a realizar los gráficos de control y meta-análisis. En las figuras 15.4 y 15.5 se muestran los resultados de los gráficos de control y meta-análisis sin *upm02*.

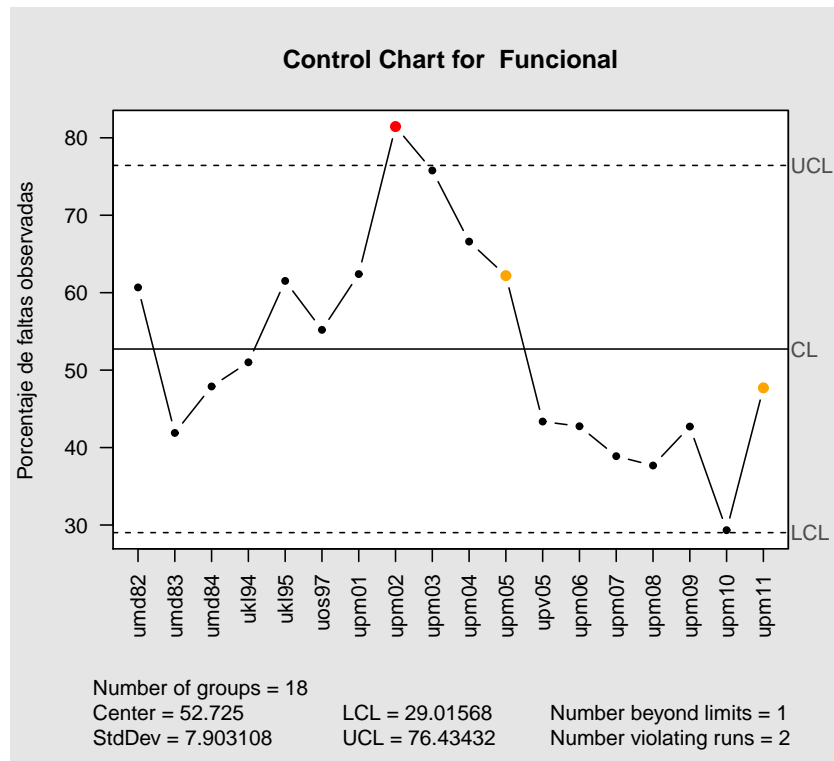
Tras apartar *upm02*, se observa que en los gráficos de control (figuras 15.4a y 15.4b) la replicación *upm03* se encuentra fuera de los límites de control. Respecto al meta-análisis de la figura 15.5 la prueba de heterogeneidad no es significativa ($Q = 21.6$, $p = 0.1565$) sin embargo el índice I^2 indica cierto porcentaje de variabilidad del 25.9%, el índice T^2 indica una mínima varianza inter-estudios ($T^2=0.0206$). Para obtener resultados consistentes se aparta *upm03* y se vuelven a realizar los gráficos de control y meta-análisis. En las figuras 15.6 y 15.7 se muestran los resultados de los gráficos de control y meta-análisis tras apartar *upm03*.

Al realizar de nuevo los análisis, los resultados de los gráficos de control se encuentran dentro de los límites de control. Sin embargo en ambos gráficos se observa una alerta que indica una posible variación en el proceso. Esta alerta se relaciona con las distintas reglas definidas en empresas del sector industrial para indicar si un proceso se considera o no fuera de control⁵.

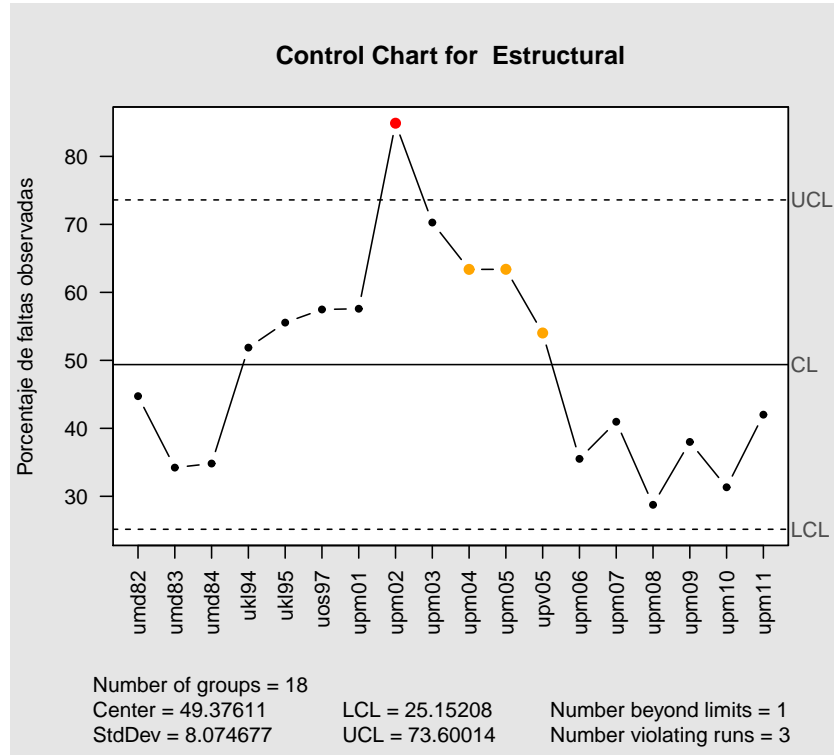
En esta primera versión del procedimiento para comparar replicaciones se usa sólo la primera regla que indica que el proceso está fuera de control cuando una o varias

⁴Gráficos generados como ficheros EPS tras cargar en el entorno R (R Dev. Core Team, 2011) el fichero `r_script02.txt`.

⁵En el apéndice B se discute este tema.



(a) Prueba funcional



(b) Prueba estructural

Fig. 15.2: Gráficos de control resultantes

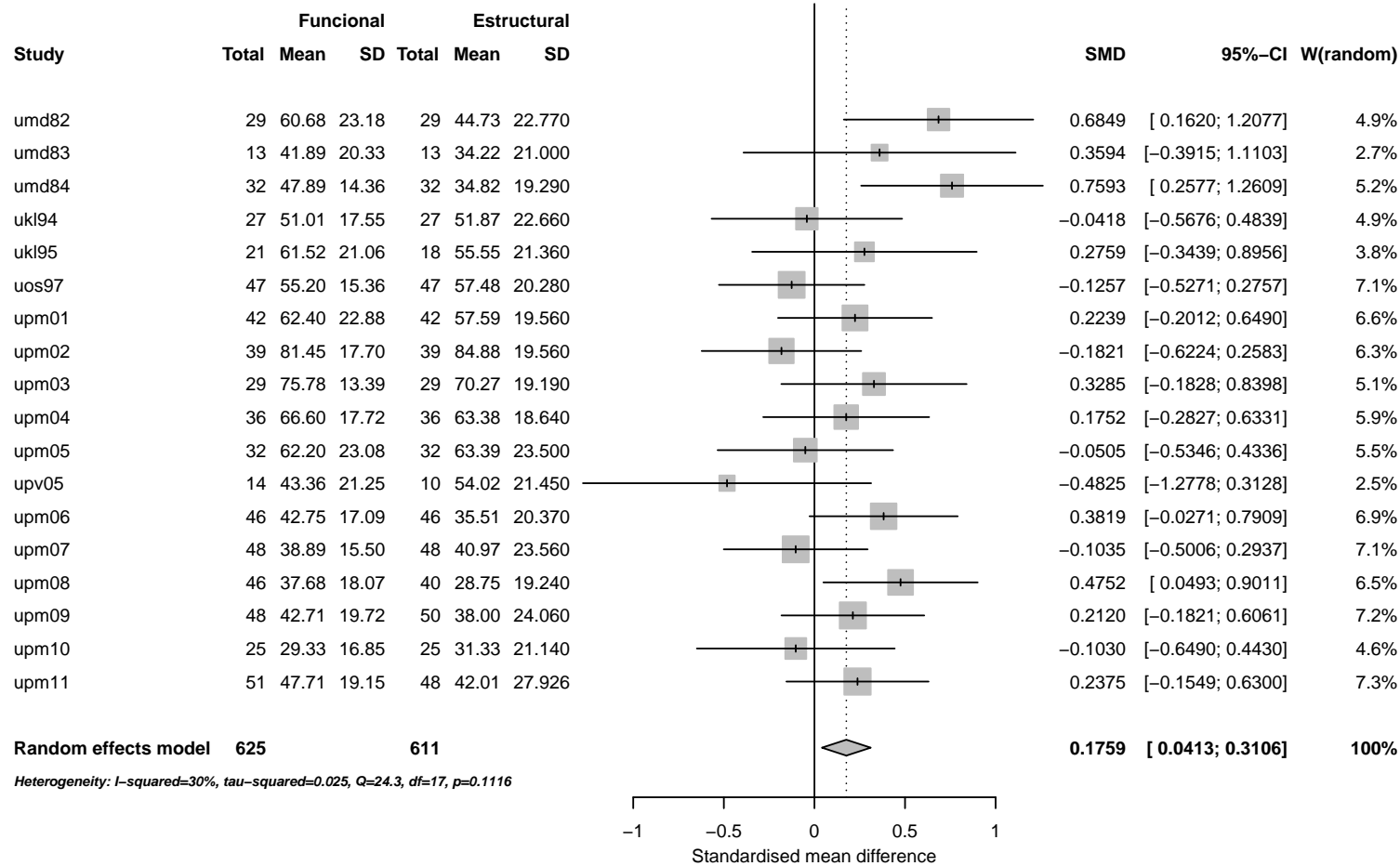
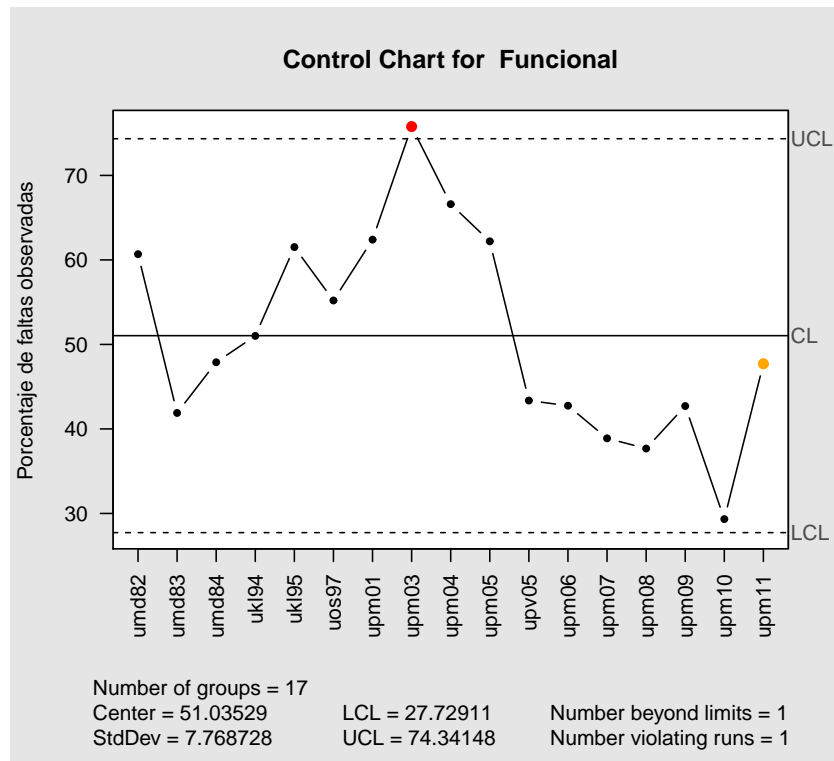
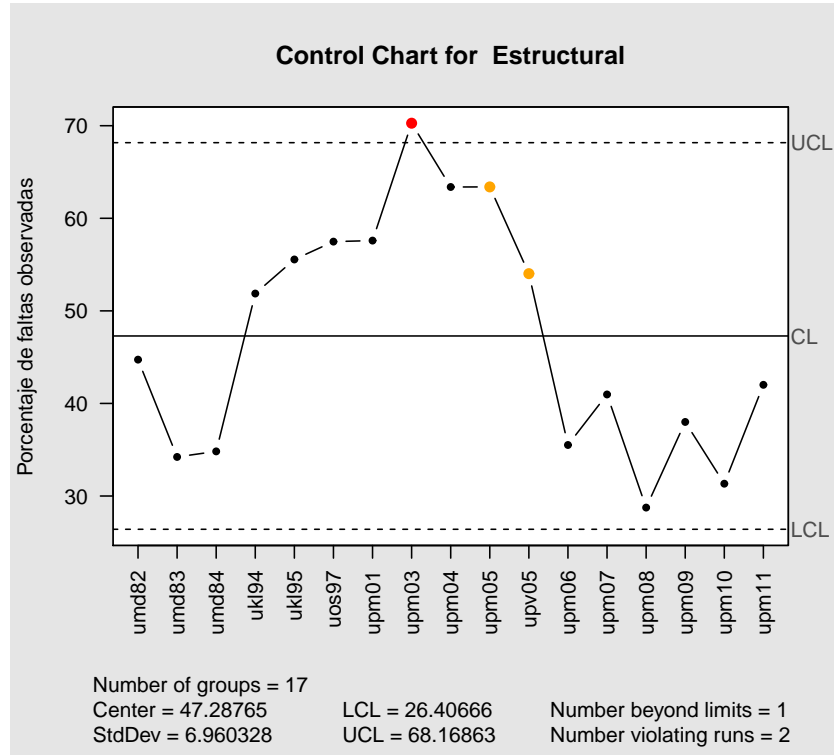


Fig. 15.3: Meta-análisis de las pruebas funcional y estructural



(a) Prueba funcional



(b) Prueba estructural

Fig. 15.4: Gráficos de control sin upm02

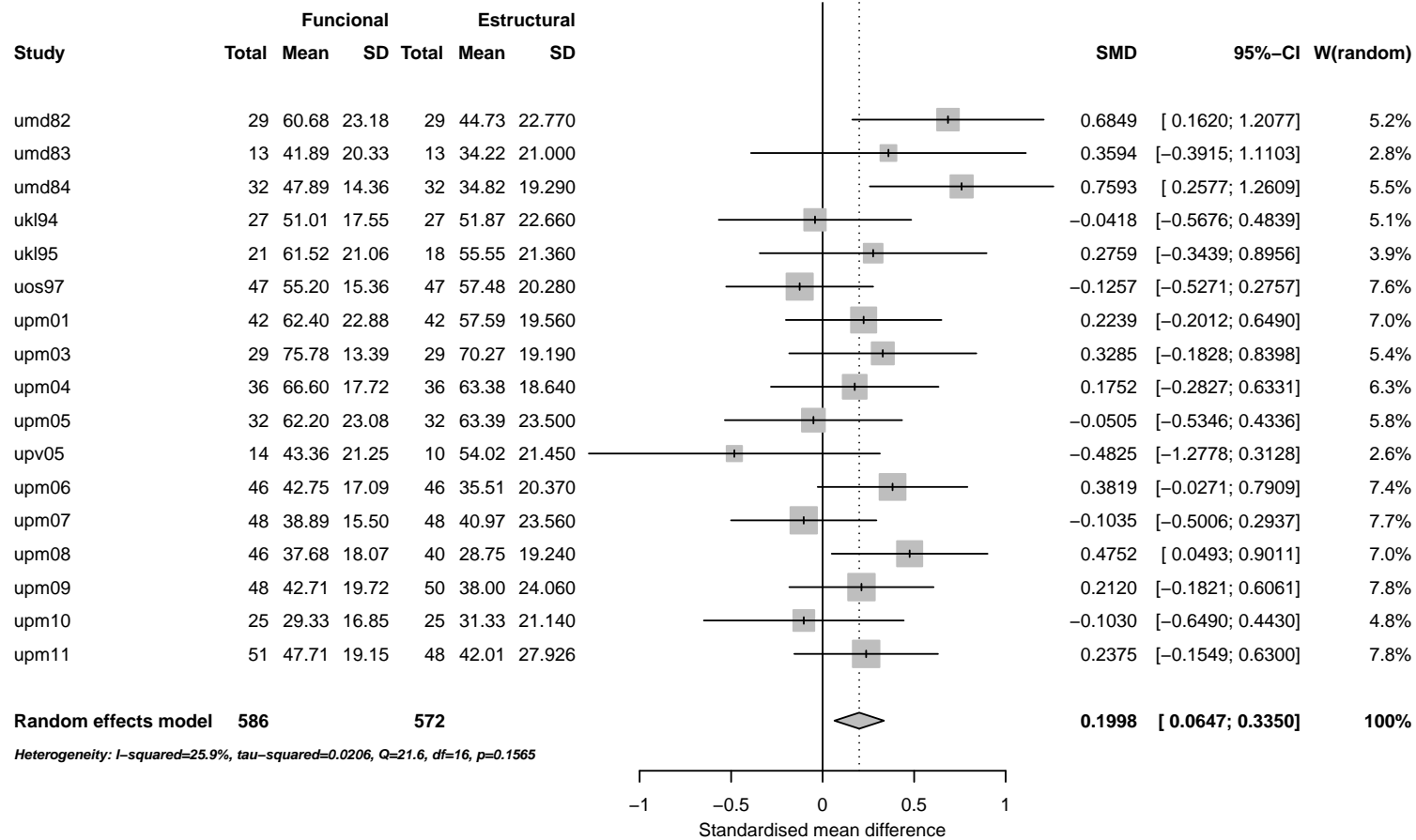
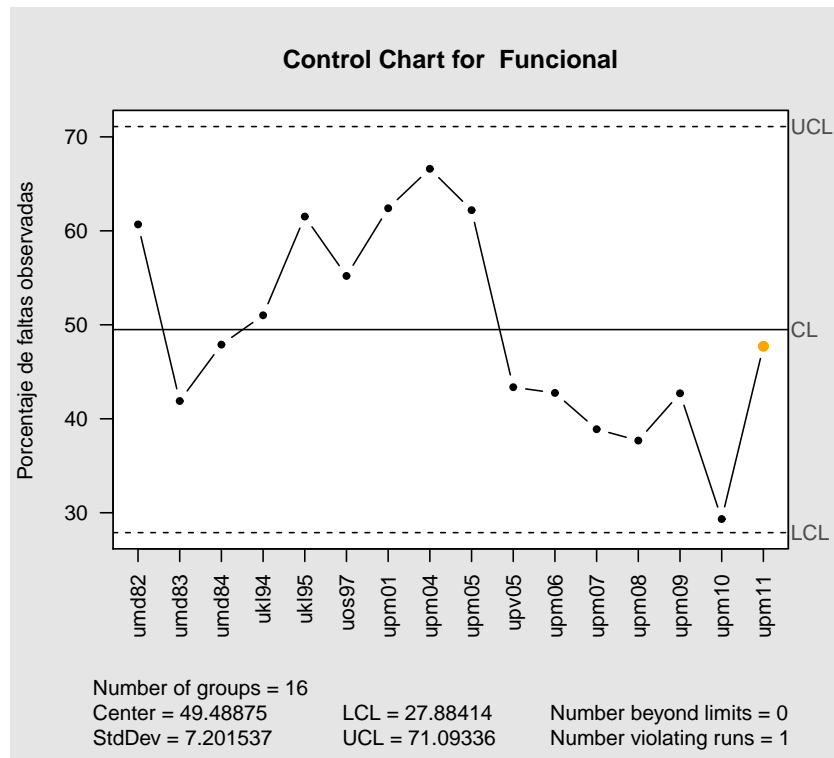
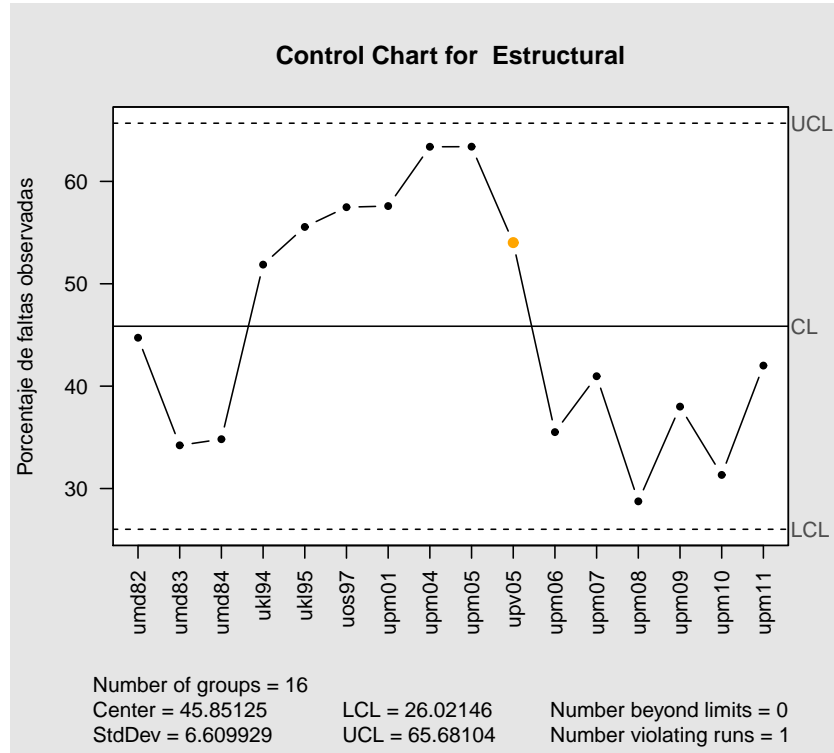


Fig. 15.5: Meta-análisis de las pruebas funcional y estructural sin upm02



(a) Prueba funcional



(b) Prueba estructural

Fig. 15.6: Gráficos de control sin upm02 y upm03

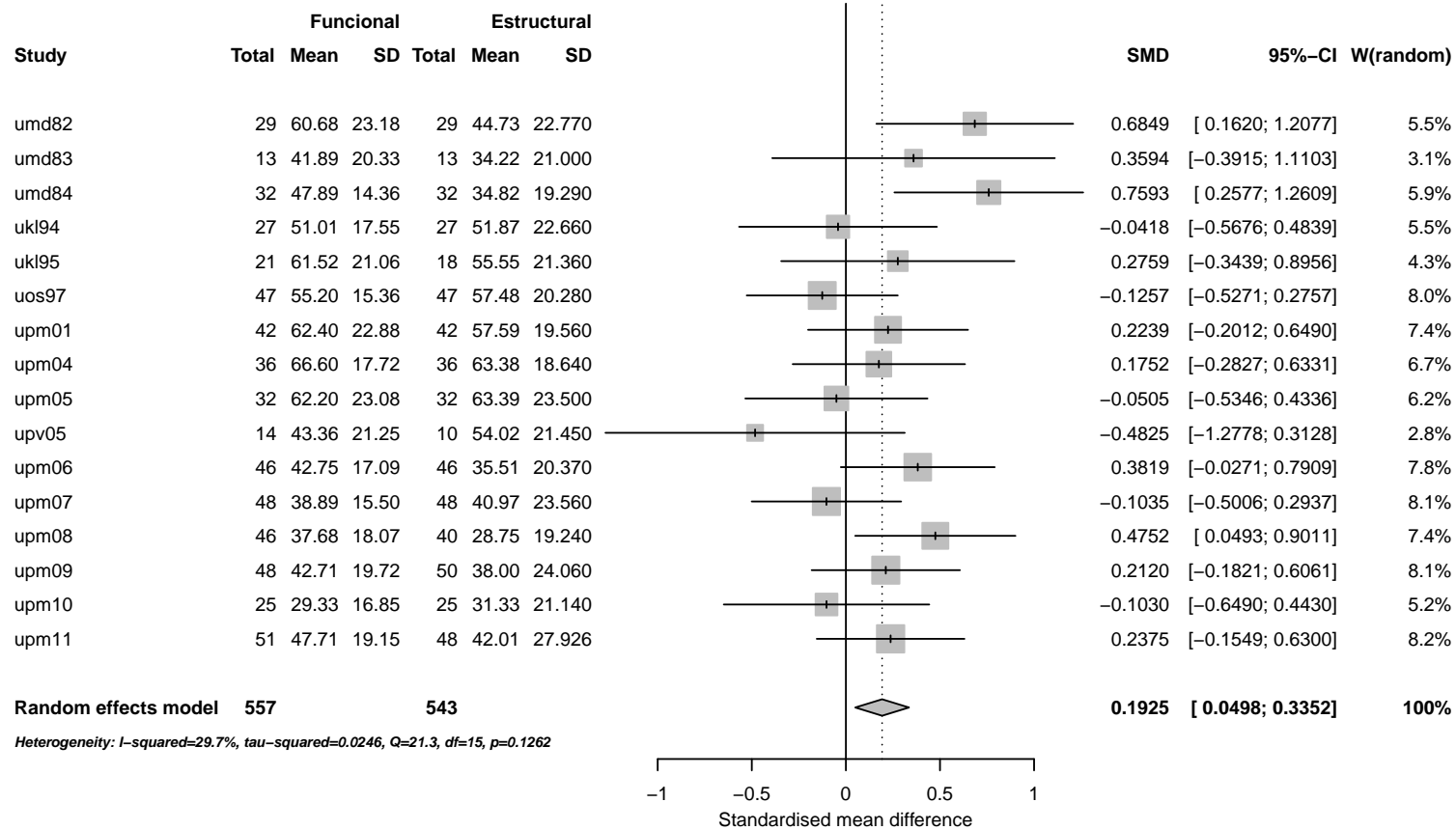


Fig. 15.7: Meta-análisis de las pruebas funcional y estructural sin upm02 y upm03

observaciones se encuentran fuera de la tercera desviación. Respecto a los distintos grupos de reglas definidas en empresas del sector industrial, hay común acuerdo en esta primera regla, el resto de grupos de reglas varía de un autor a otro. En ISE faltaría definir reglas similares de acuerdo a las observaciones recabadas de diferentes conjuntos de replications.

Respecto a los resultados del meta-análisis presentados en la figura 15.7, los resultados de la prueba de heterogeneidad no son significativos ($Q = 21.3$, $p = 0.1262$), sin embargo se observa cierto porcentaje de variabilidad ($I^2 = 29.7\%$) y una mínima varianza inter-estudios ($T^2 = 0.0246$). La replicación que parece influir en esta heterogeneidad es *umd84*. Esta replicación se aparta del resto y se vuelven a generar los gráficos de control y meta-análisis. En las figuras 15.8 y 15.9 se muestran los resultados de los gráficos de control y meta-análisis tras apartar *umd84*.

Al apartar *umd84* y realizar de nuevo los análisis, los resultados de los gráficos de control que representan las pruebas funcional y estructural (figuras 15.8a y 15.8b) se encuentran dentro de los límites control. Respecto a los resultados del meta-análisis (figura 15.9), la prueba de heterogeneidad no es significativa ($Q = 16.1$, $p = 0.3046$). Por otra parte los índices I^2 y T^2 indican un mínimo porcentaje de variabilidad ($I^2 = 13.3$) y una mínima varianza inter-estudios $T^2 = 0.0089$ respectivamente. Estos nuevos resultados se consideran consistentes para determinar el grado de verificación logrado. Sin embargo, al apartar replications en esta partición se recomienda realizar el análisis de causas para tratar de identificar los elementos de la situación experimental que se cree influyen en los resultados experimentales.

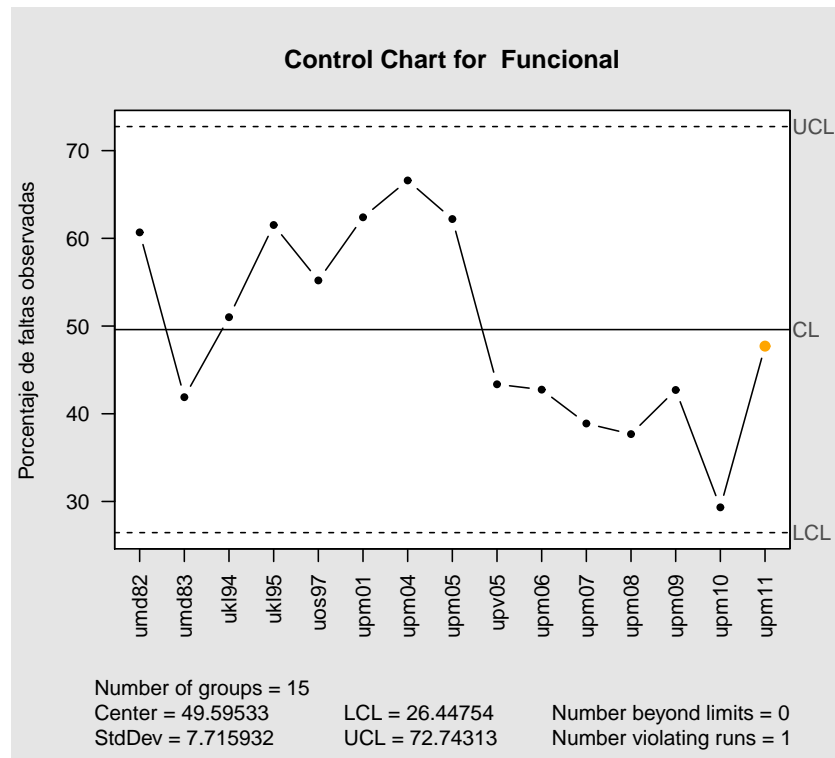
15.2.2.1 Análisis de Causas

En las figuras 15.10, 15.11, 15.12 y 15.13 se muestran los diagramas de Ishikawa (Ishikawa, 1990) con el número de cambios identificados en los distintos elementos de la situación experimental.

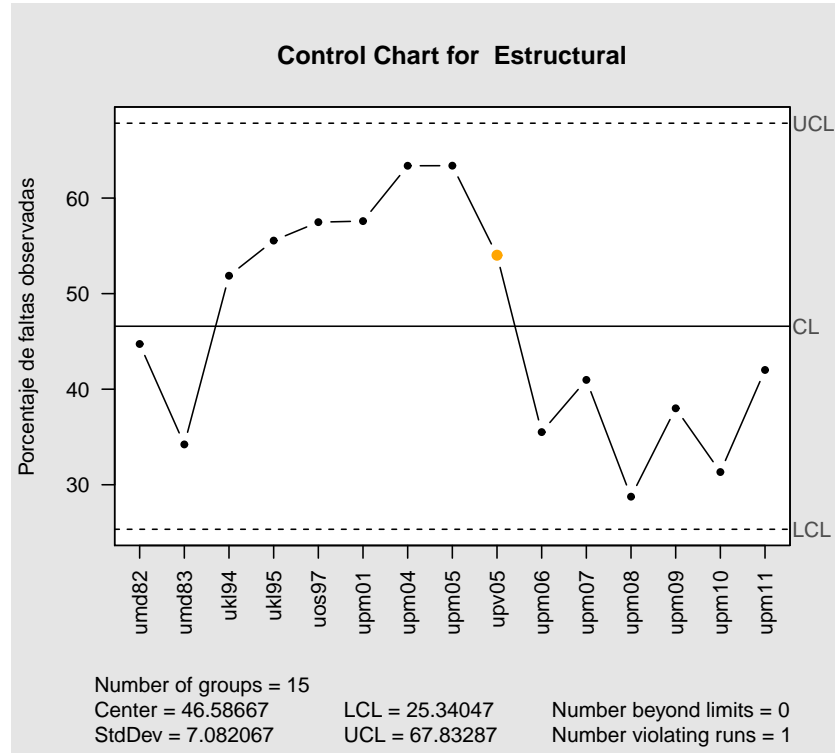
Como se observa en las figuras 15.10, 15.11, 15.12 y 15.13 la mayoría de los elementos de la situación experimental se han variado en las replications que conforman esta partición. Al haber tantos cambios en estas replications es complicado identificar los elementos que se cree influyen en los resultados de las replications *upm02*, *upm03* y *umd84*.

Tomando como referencia el análisis de causas realizado en la partición 06.b (sección 12.2.2.3) en *upm02* los cambios hechos en el rol de monitor parecen influir en los resultados. Esta replicación es más similar a *upm03*. A diferencia de *upm03*, en *upm02* participaron otros dos investigadores en el rol de monitores. Tal vez estos investigadores dieron algunas instrucciones a los sujetos que influyeron en los resultados de la replicación. Respecto a *upm03*, ésta es igual a la replicación *upm04* sin embargo la efectividad en ambas pruebas de la replicación *upm03* es mayor al resto de replications *upm* que se le parecen. En esta replicación otra condición no identificada pudo influir en los resultados. En *umd84* se cree que el tipo de sujeto influye en los resultados, en esta replicación a diferencia del resto participaron profesionales con experiencia en el uso de las pruebas funcional y estructural.

Al igual que en los análisis realizados en las particiones 11c y 09 donde las faltas parecen influir en los dos grupos de resultados identificados, se cree que en esta partición las faltas también influyen en los resultados. En la mayoría de estas replications (*umd82*, *umd83*, *umd84*, *ukl94*, *ukl95*, *uos97*, *upm01..upm05* así como *upv05*, y *upm06..upm11*) se han usado distintas instancias y tipos de faltas. Las distintas instancias y tipos de faltas



(a) Prueba funcional



(b) Prueba estructural

Fig. 15.8: Gráficos de control resultantes

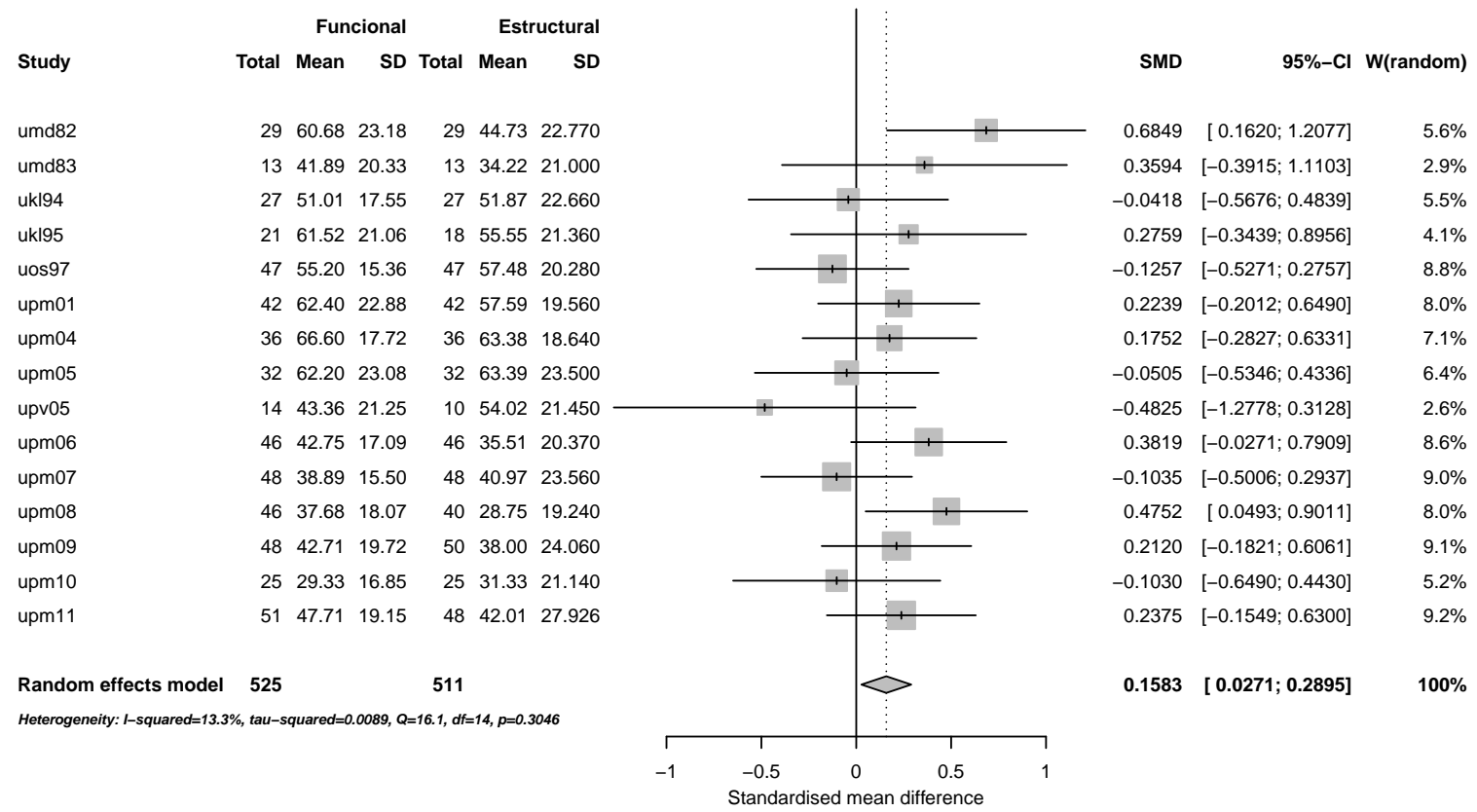


Fig. 15.9: Meta-análisis de las pruebas funcional y estructural sin umd84, upm02 y upm03

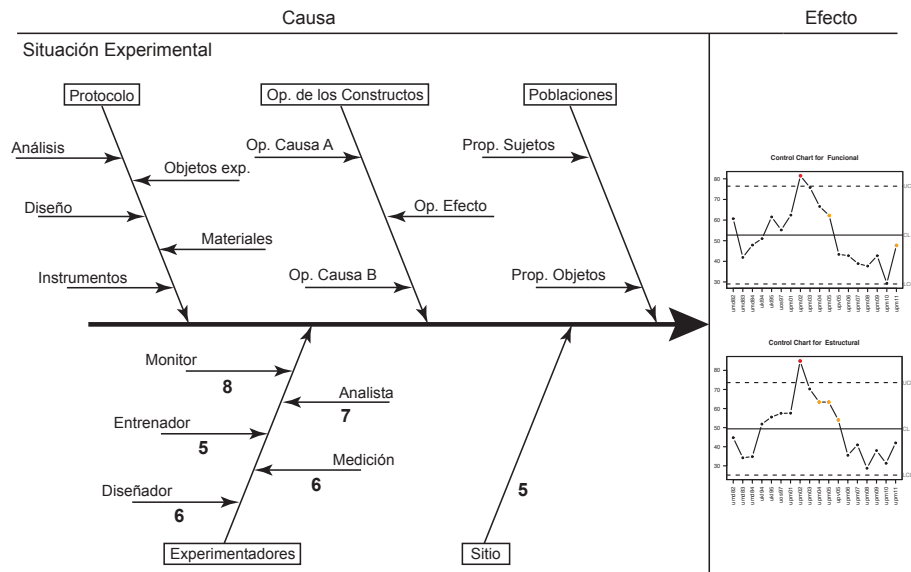


Fig. 15.10: Cambios identificados en experimentadores y sitio

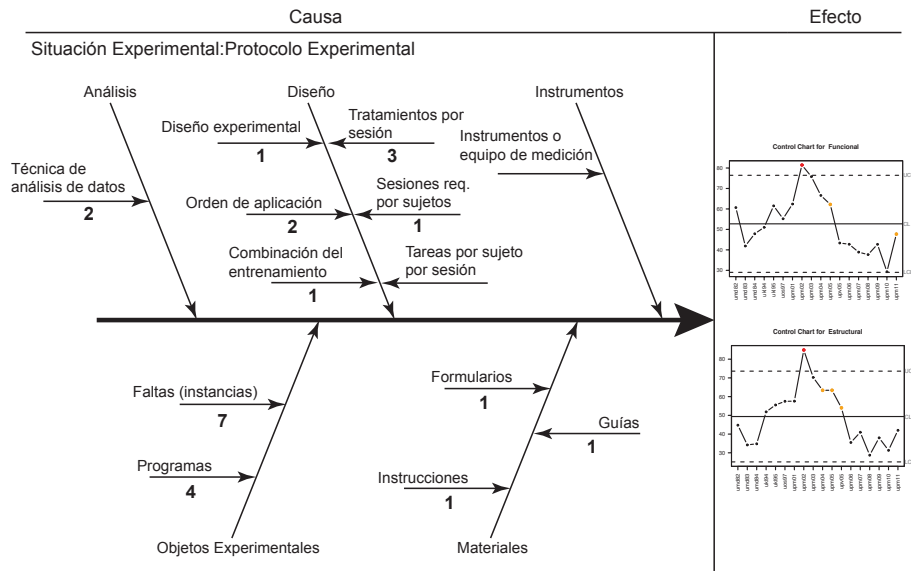


Fig. 15.11: Cambios identificados en el protocolo experimental

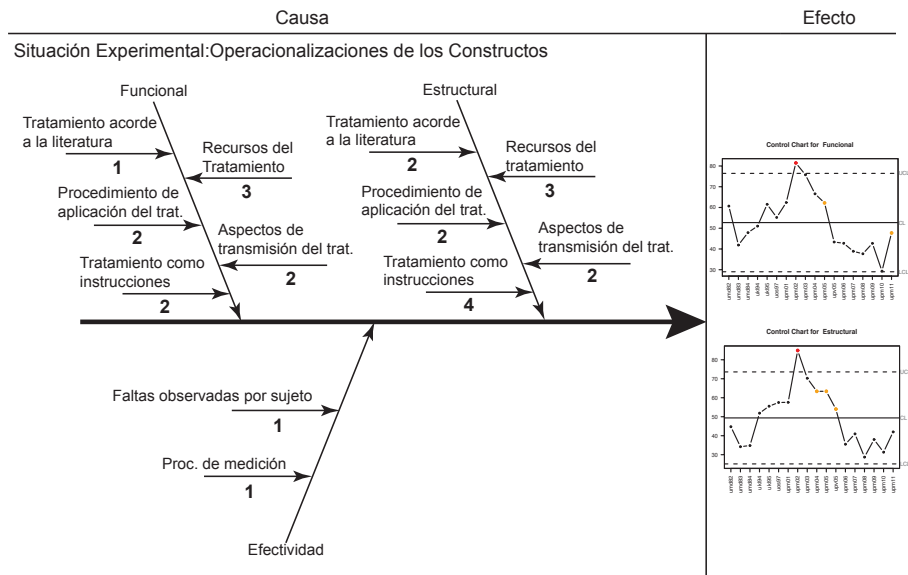


Fig. 15.12: Cambios identificados en operacionalizaciones de los constructos

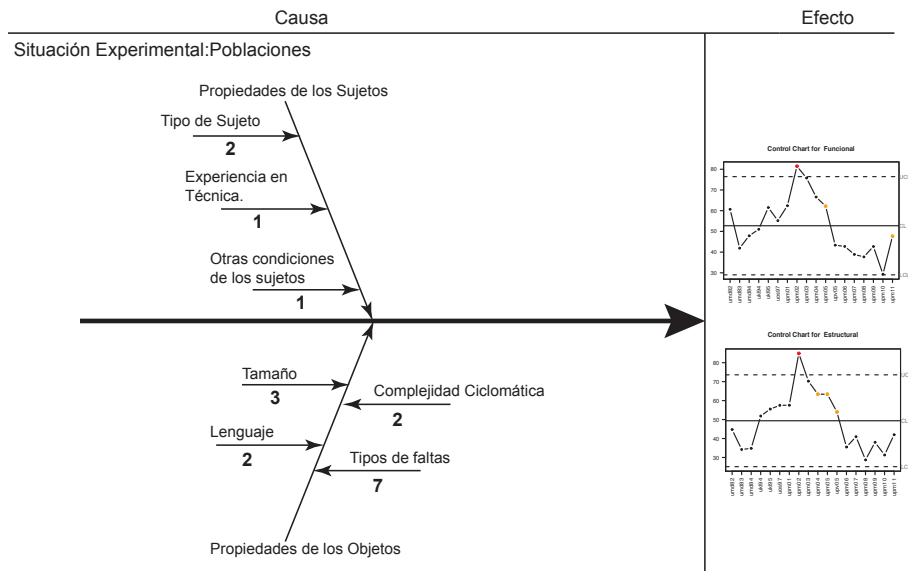


Fig. 15.13: Cambios identificados en poblaciones

parecen reflejarse en cierta medida en los gráficos de control. Por ejemplo en el gráfico de control de la prueba estructural (figura 15.8b) se observa la agrupación de las replicas *upm06* a *upm11* que usan las mismas instancias y tipos de falta.

En este gráfico se observa que los resultados de las replicas *ukl94*, *ukl95*, *uos97*, *upm01*, *upm04*, *upm05* y *upv05* se asemejan en cierta medida. Las instancias y tipos de falta usadas en estas replicas son más semejantes entre sí, en *uos97* se usaron los mismos materiales que en *ukl94* y *ukl95* sólo se modificaron y añadieron unas cuantas instancias y tipos de faltas, de manera similar las replicas *upm01* a *upm05* así como *upv05* se basan en los materiales usados en *ukl94* y *ukl95*. Estas instancias y tipos de falta difieren en mayor medida de las instancias y tipos de falta usadas en las replicas *umd82* a *umd84* y *upm06* a *upm11*.

De los cambios identificados se definen las siguientes hipótesis que pueden ser testadas en replicas posteriores:

- El rol de monitor se cree influye en los resultados experimentales.
- Las instancias y tipos de falta se cree influyen en los resultados experimentales.

Ya completada esta actividad del sub-proceso *análisis de resultados experimentales* se selecciona el grupo de resultados consistentes para conocer el grado de verificación alcanzado y empaquetar este conocimiento tanto para el profesional como para el investigador. A continuación se desarrolla la última actividad de este sub-proceso.

15.2.3 Generar Tabla con Información del Análisis Realizado

Una vez realizada la comparación de resultados experimentales, en esta actividad se coloca en el formulario PCR-9 (apéndice D.9) la información de las comparaciones de resultados realizadas en la actividad anterior. En la tabla 15.3 se muestra la estructura principal del formulario PCR-9 con la información de los resultados de las comparaciones realizadas en la actividad anterior. El grupo de resultados consistentes se ha identificado como la partición 48.1.

15.3 Sub-proceso Empaquetado de Piezas de Conocimiento

Una vez realizadas las actividades de los dos sub-procesos anteriores, a continuación se llevan a cabo las actividades del último sub-proceso del *comparador de replicas* para empaquetar el conocimiento obtenido. Este sub-proceso se compone de las siguientes actividades:

- Generar tabla de variaciones en situación experimental
- Generar pieza de conocimiento para profesional
- Generar pieza de conocimiento para investigador

Información de la Partición	
Id. Partición:	48.1
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruct.
Replicaciones de la partición:	18 (umd82..umd84, ukl94, ukl95, uos97, upm01..upm11, upv05)
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes:	Con Cambios
Tipo de resultados encontrados:	Inconsistentes (variabilidad en gráficos de control)
Información de los resultados:	<p>Q=24.3, p=0.1116, I²=30%, T²= 0.025</p> <p>Func.: 52.72% [29.01; 74.43]</p> <p>f. observadas: 60.68, 41.89, 47.89, 51.01, 61.52, 55.2, 62.4, 81.45, 75.78, 66.6, 62.2, 43.36, 42.75, 38.89, 37.68, 42.71, 29.33, 47.71</p> <p>Estruct.: 49.38% [25.15; 73.6]</p> <p>f. observadas: 44.73, 34.22, 34.82, 51.87, 55.55, 57.48, 57.59, 84.88, 70.27, 63.38, 63.39, 54.02, 35.51, 40.97, 28.75, 38, 31.33, 42.01</p>
Tipo de inconsistencia observada:	inconsistencias en gráficos de control
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento:	15 (umd82, umd83, ukl94, ukl95, uos97, upm01, upm04..upm11, upv05)
Comparación del par de tratamientos:	Efectividad en Func. ligeramente mejor a Estruct.
Tipo de efecto global:	No significativo (Usando la clasificación de Cohen el efecto es menor a 0.20)
Tamaño del efecto global:	0.1583 [0.0271; 0.2895]
Análisis de Causas	
Nuevas hipótesis de trabajo:	<ol style="list-style-type: none"> 1. El rol de monitor se cree influye en los resultados experimentales 2. Las instancias y tipos de falta se cree influyen en los resultados experimentales

Tabla 15.3: Información de las comparaciones de resultados

15.3.1 Generar Tabla de Variaciones en Situación Experimental

En esta actividad se utiliza la información de las tablas de particiones generadas en el sub-proceso *identificación de similitudes* para calcular los cambios realizados en la situación experimental de las replicaciones que constituyen los resultados consistentes de esta partición. En esta actividad se usa el formulario PCR-10 (apéndice D.10) para colocar los cálculos de los cambios realizados. La tabla resultante se utiliza en las siguientes actividades del sub-proceso para empaquetar el conocimiento obtenido. En la tabla 15.4 se muestra la estructura principal del formulario PCR-10 con la el cálculo de variaciones del grupos de resultados consistentes.

15.3.2 Generar Pieza de Conocimiento para Profesional

A partir de las tablas con información de los análisis realizados en ambos grupos de resultados (sección 15.2.3) y con las tablas de variaciones (sección 15.3.1) se empaqueta la pieza de conocimiento para el profesional. Para construir la pieza se utiliza el formulario PCR-11 (apéndice D.11). En la tabla 15.5 se muestra la estructura principal del formulario PCR-11 con la información de la pieza de conocimiento resultante para el profesional.

Información de la partición					
Id Partición:	48.1				
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruc.				
Replicaciones consideradas para la pieza de conocimiento:	15 (umd82, umd83, ukl94, ukl95, uos97, upm01, upm04..upm11, upv05)				
Variaciones en la situación experimental					
Secciones de la situación experimental	Elemento	(sub) elementos variados	total de elementos	% de variación	% variación por sección de la situación experimental
Experimentadores	Diseñador	1	1	100%	
	Entrenador	1	1	100%	
	Monitor	1	1	100%	
	Medición	1	1	100%	
	Analista	1	1	100%	100%
Sitio		1	1	100%	100%
Protocolo experimental	Análisis	1	1	100%	
	Diseño	5	6	83%	
	Instrumentos	0	1	0%	
	Objetos experimentales	2	2	100%	
	Materiales	3	3	100%	77%
Operacionalizaciones de los constructos	Tratamiento A	5	5	100%	
	Tratamiento B	5	5	100%	
	Efecto	2	2	100%	100%
Poblaciones	Sujetos	3	3	100%	
	Objetos experimentales	4	4	100%	100%

Tabla 15.4: Tabla de variaciones en situación experimental de la partición 48.1

De manera general la información que ofrece esta primera pieza de conocimiento (tabla 15.5) al profesional es la siguiente. De 18 replicaciones que conforman la partición seleccionada, se observaron inconsistencias en los resultados. Al apartar las replicaciones *umd82*, *upm02* y *upm03* la heterogeneidad es mínima por lo que es posible obtener el grado de verificación logrado con las replicaciones restantes. Los resultados de la comparación entre ambos tratamientos indican que la prueba funcional es ligeramente más efectiva que la prueba estructural sin embargo esta diferencia no es significativa. El grado de verificación logrado es el siguiente, este resultado es objetivo independiente de los experimentadores y global independiente del sitio. El resultado observado es probablemente real debido a que no se ha variado la mayor parte del protocolo experimental. En cuanto a las operacionalizaciones, se tiene conocimiento en la forma de aplicar los tratamientos y medir los resultados. Respecto a las poblaciones, se tiene conocimiento sobre las diferentes propiedades de las poblaciones usadas en estas replicaciones. La pieza de conocimiento es confiable al tener una baja heterogeneidad en las replicaciones consideradas en esta pieza, por último el poder de confirmación de los resultados es del 95% ya que se han variado la mayoría de los elementos de la situación experimental.

15.3.3 Generar Pieza de Conocimiento para Investigador

De manera similar a la actividad anterior, en esta última actividad se utiliza la tabla con información de los análisis realizados (sección 15.2.3) y la tabla de variaciones (sección 15.3.1). A partir de estas tablas se empaqueta la pieza de conocimiento para el investigador. Para construir la pieza de conocimiento se utiliza el formulario PCR-12 (apéndice

Información de la Partición		
Id Partición:	48.1	
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruct.	
Replicaciones de la partición:	18 (umd82..umd84, ukl94, ukl95, uos97, upm01..upm11, upv05)	
Resultados del análisis de similitudes:	Con Cambios	
Tipo de resultados encontrados:	Inconsistentes	
Tipo de inconsistencia observada:	inconsistencias en gráficos de control	
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento:	15 (umd82, umd83, ukl94, ukl95, uos97, upm01, upm04..upm11, upv05)	
Comparación:	Efectividad en Func. ligeramente mejor a Estruct.	
Tipo de efecto global:	No significativo (Usando la clasificación de Cohen el efecto es menor a 0.20)	
Tamaño del efecto global:	0.1583 [0.0271; 0.2895]	
Variaciones en la situación experimental		
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido
Experimentadores:	100%	Objetivo
Sitio:	100%	Global
Protocolo experimental:	77%	Semi-real
Operacionalizaciones de los constructos:	100%	Con conocimiento en la forma de aplicar los tratamientos y medir los resultados
Poblaciones:	100%	Con conocimiento de poblaciones equivalentes
Poder de confirmación de resultados:	95%	
Confiabilidad de la pieza de conocimiento:	Confiable	

Tabla 15.5: Pieza de conocimiento resultante para profesional

D.12). En la tabla 15.6 se muestra la estructura principal del formulario PCR-12 con la información de la pieza de conocimiento resultante para el investigador.

De acuerdo a la información de esta primera pieza de conocimiento (tabla 15.6), el investigador obtiene los siguientes resultados. De la partición seleccionada se observa una heterogeneidad en el meta-análisis debido a las replicaciones *umd82*, *upm02* y *upm03*. Al seleccionar el resto de replicaciones de la partición, los resultados son consistentes entre sí y es posible conocer el grado de verificación alcanzado. De acuerdo a la comparación de tratamientos la prueba funcional es ligeramente más efectiva que la prueba estructural, no obstante la diferencia no es significativa. El grado de verificación logrado es el siguiente, este resultado es objetivo independiente de los experimentadores y global independiente del sitio. El resultado observado es probablemente real debido a que se ha variado la mayor parte del protocolo experimental. En cuanto a las operacionalizaciones, se tiene conocimiento en la forma de aplicar los tratamientos y medir los resultados. Respecto a las poblaciones, se tiene conocimiento sobre las diferentes propiedades de las poblaciones usadas en estas replicaciones. Los tipos de replicaciones identificados son: *Nativa*, *Conjunta*, *Ajena*, *Interna*, *Externa*, *Artificial*, *Operacional* y *Poblacional*. Esta pieza de conocimiento es confiable al tener una baja heterogeneidad y el poder de confirmación de los resultados es del 95% debido a que se han variado la mayor parte de los elementos de la situación experimental.

De acuerdo al análisis de causas realizado, el rol del monitor así como las instancias

Información de la Partición			
Id Partición:	48.1		
Partición:	Efectividad/ Faltas Observadas/ Func. vs Estruct.		
Replicaciones de la partición:	18 (umd82..umd84, ukl94, ukl95, uos97, upm01..upm11, upv05)		
Resultados del análisis de similitudes:	Con Cambios		
Tipo de resultados encontrados:	Inconsistentes		
Información de los resultados:	Q=24.3, p=0.1116, I ² =30%, T ² = 0.025 Func.: 52.72% [29.01; 74.43] f. observadas: 60.68, 41.89, 47.89, 51.01, 61.52, 55.2, 62.4, 81.45, 75.78, 66.6, 62.2, 43.36, 42.75, 38.89, 37.68, 42.71, 29.33, 47.71 Estruct.: 49.38% [25.15; 73.6] f. observadas: 44.73, 34.22, 34.82, 51.87, 55.55, 57.48, 57.59, 84.88, 70.27, 63.38, 63.39, 54.02, 35.51, 40.97, 28.75, 38, 31.33, 42.01		
Tipo de inconsistencia observada:	inconsistencias en gráficos de control		
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento:	15 (umd82, umd83, ukl94, ukl95, uos97, upm01, upm04..upm11, upv05)		
Comparación:	Efectividad en Func. ligeramente mejor a Estruct.		
Tipo de efecto global:	No significativo (Usando la clasificación de Cohen		
Tamaño del efecto Global:	el efecto es menor a 0.20) 0.1583 [0.0271; 0.2895]		
Variaciones en la situación experimental			
Secciones de la situación experimental	Porcentaje de variaciones	Tipo de evento obtenido	Tipos de Replicaciones
Experimentadores:	100%	Objetivo	Nativa, Conjunta, Ajena Interna, Externa Artificial
Sitio:	100%	Global	
Protocolo experimental:	77%	Semi-real	
Operacionalizaciones de los constructos:	100%	Con conocimiento en la forma de aplicar los tratamientos y medir los resultados	Operacional
Poblaciones:	100%	Con conocimiento de poblaciones equivalentes	Poblacional
Poder de confirmación de resultados	95%		
Confiabilidad de la pieza de conocimiento:	Confiable		
Análisis de Causas			
Nuevas hipótesis de trabajo:	1. El rol de monitor se cree influye en los resultados 2. Las instancias de las faltas así como el tipo de falta se cree influyen en los resultados		

Tabla 15.6: Pieza de conocimiento resultante para investigador

de las faltas y el tipo de falta se cree influyen en los resultados. Estas condiciones pueden ser testadas en replicaciones posteriores.

15.4 Resultados de la Evaluación

La tipología de replicaciones propuesta ha facilitado la identificación de replicaciones hechas de forma diferenciada, ajena y externa elegidas para realizar comparaciones y obtener conocimiento de ellas. Aunque se han observado inconsistencias en los resultados de las replicaciones de esta partición, a través del *comparador de replicaciones* ha sido posible conocer el grado de verificación alcanzado en las replicaciones con resultados consistentes. Este sub-proceso también ha facilitado la identificación de variables moderadoras que se cree influyen en los resultados experimentales. Estas variables pueden ser estudiadas en profundidad en replicaciones posteriores.

Nótese que el estado de la práctica para llevar a cabo comparaciones entre replicaciones era que sólo se comparaban los resultados experimentales y se intentaba agregarlos a través de técnicas de meta-análisis, en caso de obtener resultados no reproducibles, se fracasaba en la obtención de consecuencias de la comparación de los resultados. El *comparador de replicaciones* permite al investigador comparar las situaciones experimentales de varias replicaciones, comparar los resultados experimentales no sólo a través del meta-análisis sino también empleando otras técnicas que facilitan examinar las variaciones de los tratamientos. Éste también facilita la identificación de variables moderadoras cuando se obtienen resultados no reproducibles. Las distintas comparaciones hechas en este sub-proceso permiten conocer el grado de verificación logrado en las replicaciones de una partición y empaquetar este conocimiento para que sea usado por el profesional o el investigador.

Parte IV

Conclusiones

Capítulo 16

Conclusiones y Futuras Líneas de Investigación

—It is as though all students of violin were told that they must avoid playing songs that had already been played. Research is not only a creative process, it is a discipline. Some concertos are best understood by being played more than once.

(Easley et al., 2000)

RESUMEN: En este capítulo se se presentan las principales contribuciones de esta investigación, los resultados obtenidos así como las futuras líneas de investigación.

La replicación desempeña un papel importante en el progreso científico, en ciencia los hechos son al menos tan importantes como las ideas (Hunter, 2001). La replicación de experimentos es necesaria para la determinación de hechos. Si se desea fortalecer el cuerpo de conocimientos que se tiene en IS es necesario llevar a cabo diferentes tipos de replications donde cada tipo permita verificar resultados experimentales de acuerdo a propósitos específicos.

Aunque existen avances en el uso de experimentos en IS, es necesario contar con más replications para comprender mejor los mecanismos que influyen en la construcción del software y así consolidar un cuerpo de conocimientos en IS que pueda usarse tanto por profesionales como por investigadores, en palabras de (Basili et al., 1999) “para construir un cuerpo de conocimientos se requiere de familias de experimentos así como un conjunto de principios unificados que permitan combinar y generalizar resultados”, o como afirman Sjøberg et al. (Sjøberg et al., 2007) “para la construcción de más y mejores teorías en IS, la comunidad necesita llevar a cabo estudios empíricos que sean relevantes y de alta calidad, incluyendo combinaciones extensivas de estudios en la forma de familias de replications, así como sintetizar la evidencia de tales estudios”.

De manera general, en esta tesis se ha trabajado en la adaptación del concepto de replicación a la ISE, específicamente se ha trabajado en 1) clarificar el concepto de replicación en ISE, 2) desarrollar una tipología de replications para la ISE y 3) desarrollar un marco conceptual con ideas clave para la comparación de replications. Las ideas de este

marco implican entender cuáles son las unidades comparables entre replicaciones; comparar gran cantidad de información referente a los elementos de la situación experimental de cada replicación; comparar resultados experimentales; combinar toda esta información para decidir si se han verificado los resultados y en qué grado; y finalmente empaquetar el conocimiento obtenido.

16.1 Contribuciones de la Investigación

La presente investigación aporta las siguientes contribuciones habiendo logrado la comprensión de cuestiones que, hasta ahora, eran confusas o desconocidas en ISE:

1. **Clarificación del concepto** de replicación a la ISE. En esta investigación se han clarificado:
 - Las diferencias entre repetición, replicación y reproducción
 - Utilidad de la repetición en ISE
 - El papel de la replicación independiente
 - Las variaciones en la replicación. Una replicación puede tener tanto los mismos elementos de la situación experimental del experimento base como puede variar uno o varios de sus elementos i.e. en la replicación es válido tanto reutilizar el protocolo del experimento base como variarlo, realizar la replicación de forma independiente o de forma conjunta.
 - Las funciones de la replicación. Cada forma de replicar un experimento satisface una función particular en la verificación de resultados experimentales.
2. Desarrollo de una **tipología de replicaciones** para la ISE. Los diferentes tipos de replicaciones encontrados en otras disciplinas describen cambios o variaciones en la estructura del experimento a ser replicado. Esto significa que una replicación puede tener diferentes niveles de similitud respecto al experimento base. Los cambios que una replicación puede tener se relacionan con las diferentes funciones de la replicación. Aunque la función general de la replicación es verificar un resultado experimental, la replicación tiene funciones específicas dependiendo de los cambios que se realicen en los elementos de la situación experimental.

La tipología de replicaciones desarrollada en esta investigación tiene como finalidad orientar al investigador en IS sobre los diferentes tipos de replicaciones que puede realizar. Esta tipología ayuda a los investigadores a planificar y entender la contribución que realizarán las replicaciones que lleven a cabo. Dependiendo de los cambios que se deseen o necesiten realizar respecto al experimento base, los investigadores pueden elegir distintos tipos de replicaciones que implican diferentes grados de verificación de resultados experimentales.

La tipología aquí propuesta contribuye a ubicar en un marco más general las visiones opuestas sobre la replicación que tienen los diferentes investigadores en ISE. A tenor de lo encontrado en esta investigación, parece que tales visiones contrarias, en realidad hablan de diferentes tipos de replicación con diferentes propósitos. La tipología propuesta cubre las distintas formas de replicación que permiten avanzar de forma gradual y sistemática hacia la verificación de resultados experimentales.

3. Desarrollo de un **marco conceptual** con ideas clave para la comparación de replicaciones en ISE. Este marco permite la explotación de conjuntos de replicaciones

con el fin de identificar el grado de verificación alcanzado y empaquetar este conocimiento en piezas que sean útiles tanto para el profesional como para el investigador. Este marco se compone de las siguientes ideas clave:

- Unidades comparables entre replicaciones. Para poder comparar conjuntos de replicaciones y obtener conocimiento de ellas, es necesario identificar las unidades comparables que tienen en común, donde cada parte se compone por un constructo y operacionalización de efecto y un par de constructos y operacionalizaciones de causa. Las agrupaciones de unidades comparables reciben el nombre de partición.
 - Cambios y nivel de similitud entre situaciones experimentales. Para conocer el grado de verificación alcanzado en un conjunto de replicaciones con resultados reproducidos, es necesario identificar los cambios hechos en los elementos de la situación experimental de las replicaciones. Dependiendo de los elementos que varíen en la situación experimental se obtiene un grado de verificación determinado. Las similitudes entre las replicaciones ofrecen una idea sobre la cantidad de cambios que hay en las situaciones experimentales de las replicaciones conformadas en una partición.
 - Variabilidad de los tratamientos. No basta con usar sólo técnicas de agregación de resultados para combinar replicaciones. Respecto a esta idea clave, se propone previo al meta-análisis examinar la variabilidad de los tratamientos de las replicaciones conformadas en una partición, ya que puede presentarse la situación donde se observe homogeneidad de resultados cuando en realidad hay variabilidad considerable en alguno de los tratamientos. En esta situación se puede concluir erróneamente que los resultados son verificables debido a que sólo se empleó meta-análisis.
 - Grado de verificación logrado. Con los cambios identificados en las situaciones experimentales más las comparaciones de los resultados experimentales es posible conocer el grado de verificación logrado en una partición cuando se obtienen resultados reproducibles. El grado de verificación permite conocer al investigador el estado de los resultados que se ha reproducido. Si éstos son locales producto de las replicaciones hechas en un mismo sitio u obedecen a una realidad debido a que los resultados se han reproducido en distintos sitios y variando el protocolo experimental. O si los resultados son generalizables a determinados tipos de sujetos o artefactos. Este conocimiento obtenido puede empaquetarse en una o varias piezas que sean útiles para el profesional y el investigador.
 - Identificación de posibles causas de resultados no reproducibles. Cuando se presentan resultados no reproducibles en una partición, es posible llevar a cabo un análisis de causas para identificar posibles elementos de la situación experimental que se cree influyen en los resultados. Los elementos identificados sirven para definir nuevas hipótesis y testarlas en replicaciones posteriores.
4. Siguiendo las ideas de este marco conceptual se ha desarrollado un **procedimiento para comparar replicaciones**. Este procedimiento visto en forma de proceso se compone de dos sub-procesos, *generador de particiones* y *comparador de replicaciones*. Para obtener conocimiento de un conjunto de replicaciones, primero es necesario identificar las unidades comparables comunes entre las replicaciones. El *generador de particiones* recibe como entrada un conjunto de replicaciones para

dividirlo en subconjuntos de replicaciones de acuerdo a los constructos y operacionalizaciones que tienen en común. Estos subconjuntos denominados como *particiones* son usados por el siguiente sub-proceso (*comparador de replicaciones*). El *comparador de replicaciones* compara tanto las situaciones experimentales de las replicaciones conformadas en una partición así como compara los resultados experimentales. Los resultados de estas comparaciones permiten conocer el grado de verificación logrado así como empaquetar este conocimiento.

5. El conjunto de replicaciones usado para la evaluación consta de una serie de experimentos replicados basados en el experimento realizado por Basili y Selby (Basili y Selby, 1987) donde se evalúa la efectividad, eficiencia y coste de tres técnicas de evaluación de software, específicamente prueba funcional (caja negra) utilizando particiones de clases de equivalencia y análisis de valores límite, prueba estructural (caja blanca) utilizando cobertura de sentencias y lectura de código por abstracciones sucesivas.

Parte de las replicaciones usadas en esta evaluación han sufrido intentos fallidos de combinar los resultados experimentales usando técnicas clásicas de agregación de resultados, tanto estadísticas (Miller, 2000) como narrativas (Runeson et al., 2006). Debido a las diferencias en los resultados de las replicaciones, en ambos casos los autores concluyen que las replicaciones son difíciles de agregar y por tanto es difícil obtener consecuencias de ellas.

Al aplicar la solución propuesta en esta tesis sobre este conjunto de replicaciones, se ha logrado satisfactoriamente combinar resultados. Ha sido posible identificar las unidades comparables comunes en estas replicaciones y agruparlas en particiones. Seguramente el fracaso de (Miller, 2000) y (Runeson et al., 2006) provenga del intento de comparar replicaciones de manera completa y no por sus partes comunes, lo que resulta inviable.

Del conjunto de particiones seleccionadas, a diferencia de las conclusiones descritas en (Miller, 2000) y (Runeson et al., 2006), ha sido posible decidir sobre el grado de verificación logrado en las distintas particiones y generar así conocimiento útil para el profesional y el investigador.

Durante la evaluación se observó que las replicaciones utilizadas se realizaron sin ningún objetivo de verificación concreto. Como es habitual en ISE, la verificación de resultados no era bien entendida y se entendía como un único objetivo. Las replicaciones presentan a la vez varios cambios en los elementos de la situación experimental (seguramente por la ignorancia del contexto). Cuando en una replicación se varía a la vez varios elementos de la situación experimental y los resultados respecto al experimento base son contradictorios es más difícil identificar la condición causante de la variación de resultados. Aunque en este conjunto se tienen replicaciones con varios cambios hechos a la vez y con resultados inconsistentes, se han logrado identificar posibles condiciones que se cree influyen en varios resultados experimentales de este conjunto de replicaciones.

6. Esta tesis propone acometer los distintos tipos de replicación ordenadamente lo cual debe permitir verificar resultados experimentales de manera gradual, aunque dando cabida a los inevitables cambios que, hoy en día, suele exigir la variación de lugar para un experimento en IS. En el estado actual de la ISE, la forma más apropiada de realizar replicaciones es a través de un **enfoque sistemático** que permita avanzar de forma gradual en el descubrimiento de nuevas condiciones que influyen en los

resultados de los experimentos (y, por tanto, en la construcción de software). Al contar con un mayor entendimiento de estas condiciones, nuestra disciplina será capaz de consolidar, a partir de pequeñas piezas (i.e. replicasiones variadas), un cuerpo de conocimientos que ayude al profesional en IS a tomar decisiones mucho más objetivas sobre algún aspecto de la construcción de software o sobre alguna tecnología software nueva por adoptar.

16.2 Publicaciones

A continuación se listan los resultados de esta investigación presentados en diferentes sitios.

1. Omar S. Gómez. *Clasificación de Replicasiones para la Síntesis de Experimentos en Ingeniería del Software*. **XIII Congreso Iberoamericano en Software Engineering** (CIbSE'2010), Cuenca, Ecuador, Abril, Pages 319-323, 2010. ISBN 978-9978-325-10-0.

El objetivo principal de este artículo fue mostrar un panorama general de la investigación doctoral que se llevaba en curso. Se menciona desde una perspectiva general el procedimiento para comparar replicasiones así como algunas funciones de la replicación. Este artículo se relaciona con parte del estado de la cuestión descrito en el capítulo 2 y con parte de la solución descrita en los capítulos 7, 8 y 9.

2. Omar S. Gómez, Natalia Juristo, Sira Vegas. *Replication, reproduction and re-analysis: Three ways for verifying experimental findings*. **International Workshop on Replication in Empirical Software Engineering Research** (RESER'2010), Cape Town, South Africa, May, 2010.

El objetivo de este trabajo preliminar fue presentar en un taller (*workshop*) específico sobre el tema de replicación para recoger comentarios sobre la investigación en curso. El trabajo recibió una buena acogida por parte de los asistentes. Este artículo presenta la contribución donde se clarifican los conceptos de replicación, reproducción y re-análisis así como se discute el tipo de verificación que cada uno conlleva. Esta contribución se describe en el capítulo 4 de esta tesis.

3. Omar S. Gómez, N. Juristo, Sira Vegas. *Replication types in experimental disciplines*. **4th International Symposium on Empirical Software Engineering and Measurement** (ESEM'2010), Bolzan, Italy, September, 2010. ISBN 978-1-4503-0039-1.

En este trabajo se presentaron algunos resultados sobre las diferentes formas de replicación empleadas en otras disciplinas científicas. Este trabajo se presentó en el congreso internacional para temas de experimentación en IS. Este trabajo se relaciona con el estado de la cuestión descrito en el capítulo 2 y con la contribución de esta tesis relacionada con la tipología de replicasiones para ISE.

4. Natalia Juristo, Omar S. Gómez. *Replication of Software Engineering Experiments*. **Empirical Software Engineering and Verification: LASER Summer School 2008-2010**, Berlin, Lecture Notes in Computer Science, Volume 7007, Pages 60-88, Springer-Verlag, 2011. ISBN 978-3-642-25230-3.

En este trabajo se discute y estudia el concepto de la replicación para obtener mayor comprensión de su aplicación a la ISE. Este trabajo se ha publicado en un capítulo

de un libro cuyo tema principal es la Ingeniería del Software Empírica. En este artículo se presenta la contribución de esta tesis sobre la necesidad de repeticiones de experimentos en ISE, así como los resultados sobre los distintos tipos de verificación que una replicación puede pretender.

5. Omar S. Gómez, Natalia Juristo, Sira Vegas. *Understanding Replication of Experiments in Software Engineering: A Classification*. Previsto enviar al IEEE Transactions on Software Engineering.

En este artículo se presenta gran parte de la investigación doctoral. Se tiene previsto enviar este artículo a una revista con arbitraje internacional especializada en IS. En este artículo se presenta la contribución de la tesis relacionada con los elementos de la situación experimental que deben tenerse en cuenta a la hora de comparar replicaciones. Los tipos de replicación que aparecen según los cambios realizados en los diferentes elementos así como su relación con los diversos grados de verificación de resultados.

6. Omar S. Gómez, Natalia Juristo, Oscar Dieste. *Repetitions as Starting Point for Replicating Software Engineering Experiments*. Previsto enviar al **6th International Symposium on Empirical Software Engineering and Measurement (ESEM'2012)**, Lund University, Sweden, September, 2012.

En este artículo se presenta la contribución de la tesis relacionada con el papel que desempeña la repetición de experimentos en la verificación de resultados experimentales.

7. Omar S. Gómez, Natalia Juristo, Sira Vegas. *Framework for Comparing Software Engineering Experiment Replications*. Previsto enviar al **Journal on Empirical Software Engineering**.

En este artículo se presenta la contribución de la tesis referente al marco conceptual para la comparación de replicaciones en ISE. En este artículo se emplea el conjunto de replicaciones usado para la evaluación de esta tesis.

Esta investigación se enmarca dentro del proyecto 206747 del Consejo Nacional de Ciencia y Tecnología CONACyT de México, así como por los proyectos TIN2008-00555 (generación de evidencias mediante combinación de resultados experimentales) y TIN2011-23216 (tecnologías para la replicación y síntesis de experimentos en IS) del Ministerio de Ciencia e Innovación del Gobierno de España.

16.3 Futuras Líneas de Investigación

De manera general en esta investigación se ha trabajado en la adaptación del concepto de replicación a la ISE, no obstante los distintos aspectos aquí analizados, así como las limitaciones encontradas, dan pie a una serie de mejoras y futuros trabajos de investigación que a continuación se detallan:

- Tomando como referencia la tipología de replicaciones aquí propuesta, sería útil realizar una investigación para identificar los diferentes tipos de replicación hechos en ISE. Se trataría de una revisión sistemática sobre replicaciones que debería identificar los cambios realizados por cada replicación con respecto al experimento base y comprobar si los tipos de la clasificación aquí propuesta permiten acoger las replicaciones encontradas.

- Los distintos cambios que se han realizado en los elementos de la situación experimental de las replications usadas para la evaluación, así como los diferentes resultados experimentales aquí presentados sirven como punto de partida para llevar a cabo futuras replications sistemáticas.

Por ejemplo, en el conjunto de replications usado no se encontró ninguna replicación literal (repetición) donde se verifique que los resultados experimentales no sean producto del error muestral. Se podría tomar como experimento base alguna de las replications más recientes y realizar varias repeticiones. Otra alternativa podría ser la siguiente, realizar dos replications donde sólo se modifiquen las propiedades de las poblaciones, por ejemplo en la primera replicación utilizar profesionales en lugar de estudiantes de pregrado, mientras que en la segunda replicación utilizar sólo estudiantes de postgrado. Los resultados de estas dos replications se contrastarían con las replications *umd82*, *umd83* y *umd84* donde se usa este tipo de sujetos. Estas replications ayudarían a verificar si el nivel de experiencia de los distintos tipos de sujetos influye en la efectividad de alguna de las técnicas de evaluación de software usadas en este conjunto de replications.

Conforme se realicen diferentes tipos de replications en este conjunto de replications, será posible contar con mayor conocimiento de las condiciones que influyen en los resultados de esta familia de experimentos. Al tener mayor diversidad de replications el conocimiento ganado será suficientemente maduro para conocer empíricamente el comportamiento de las técnicas de evaluación de software estudiadas en este conjunto de replications.

- Con el objetivo de refinar y mejorar las distintas actividades y sub-procesos que conforman el procedimiento para comparar replications faltaría llevar a cabo este proceso en otro conjunto de replications distinto al usado en esta investigación.
- En esta primera versión del procedimiento para comparar replications se requiere de cierto esfuerzo para ejecutar los sub-procesos. Con el fin de reducir dicho esfuerzo se podría construir una herramienta donde se automaticen las distintas partes que conforman este procedimiento.

Parte V

Apéndices

Apéndice A

Descriptions of the Replications Typologies

A.1 Adams et al.

Authors (Adams et al., 2005) define the following replication types: *Literal Replications*. Which are those studies where both measures and manipulations are as identical as possible to a previous study. Efforts are made to duplicate both the dependent and independent variables as closely as possible.

Operational Replications. Which are those studies that alter the operationalization of the criterion variable while maintaining the use of treatments. It is desirable to retain the conceptual integrity of the measure and the independent variables are unaltered.

Instrumental Replications. The treatments or independent variables are changed and the measures remain the same as in prior studies.

Constructive Replications. Which vary both the criterion and the manipulations or treatments. An attempt to replicate conceptually a previous study is made.

A.2 Bahr et al.

Types A..P. This classification (Bahr et al., 1983) categorizes replications according to four dichotomic properties (equal or different) of a replication. These properties are: time, place, subjects and methods. Based on combinations of these properties, Bahr et al. define 16 replication types.

A.3 Barker and Gurman

Authors (Barker y Gurman, 1972)¹ adopt the convention of viewing an experiment as consisting of two sets of variables; methodological and procedural matters on the one hand and outcome or dependent variables on the other.

Within this framework the experimenter may seek to replicate a reference experiment with respect to: 1) both the procedural and dependent variables, 2) only the dependent variables, 3) only the procedural variables, 4) neither the procedural nor dependent

¹Debido a la configuración del documento en el idioma Castellano, el comando `\citep{}` utiliza la letra ‘y’ en lugar de la palabra ‘and’ cuando la referencia consta de dos autores.

variables.

Procedural variables replicated; dependent variables replicated. The experimenter seeks to replicate both the procedural and outcome variables of a reference experiment. An attempt is made to duplicate exactly the sampling procedure, experimental conditions, measuring techniques and methods of analysis of the reference experiment.

Procedural variables varied; dependent variables replicated. In this type of replication, the experimenter seeks deliberately to vary the methodology and procedures of a reference experiment but aims also to replicate the results of the reference experiment. He tests the generality of some recognized lawful relationship.

Procedural variables replicated; dependent variables not replicated. In this type of replication, the experimenter does not use the same dependent variables as in the reference experiment.

Procedural variables not replicated; dependent variables not replicated. En esta replicación In this type of replication, the experimenter deliberately varies the methods and procedures used in the reference experiment. Such a replication and the reference experiment are equivalent in terms of the concept of the variables used in the two experiments.

A.4 Beck

Author (Beck, 1994) groups different replication classifications into one ordering, which includes the following replication types.

Type 1. Identical or literal studies are considered exact replications. *Type 2.* Virtual, operational, direct, and retest are terms used to indicate that a study is closely replicated. This strategies approximate the original research design. *Type 3.* Systematic extension, constructive, systematic, and independent replications include the extension of an earlier study. This strategies increase empirical generalization by significantly modifying the original design. *Type 4.* Pseudo-replication and Internal replication, in this strategies data from both the original and the replication studies are collected at the same time. *Type 5.* Theoretical replication, this strategy develops and verifies a theory.

A.5 Blomquist

There are two types of replication (Blomquist, 1986): *literal* and *construct*. In a *literal replication*, the researcher uses the same measures with the same type of subjects, and controls the same conditions. The original study is replicated as exactly as possible.

In a *construct replication*, a second investigation begins with a similar hypothesis but uses new methods of measurement and design. The findings support the construct under investigation and can make the findings more generalizable.

A.6 Brown and Coney

Authors define the following two replication types (Brown y Coney, 1976): *Replication*. A substantial duplication of a previously published, data based research project. Substantial means that any changes made deal with increasing the internal validity of the research effort (i.e. any change affecting the controls of a design).

Replication with extension. A duplication of a previously published, data based research project with changes made in either the active (manipulated) or attribute (measured) variables, the intent of which is to increase the external validity of the research effort (i.e. the generalizability).

A.7 Easley et al.

Authors (Easley et al., 2000) describe four replication types: *Type 0 (Precise Duplication)*. This replication is defined as a precise duplication of a prior study. Therefore, Type 0 (precise duplication) studies are those studies in which every nuance of the experimental setting is precisely reproduced; as such, the cause-effect relationship is finite. The ability to conduct a Type 0 replication is limited to experimenters in only some of the natural sciences. As others have stated, it is an impossibility to conduct a Type 0 replication in a social science context because uncontrolled extraneous factors have the potential to interact with the various components in an experimental setting. For example, human subjects cannot be precisely duplicated. A social scientist is limited only to matching subjects as closely as possible.

Type I (Duplication). A type I replication is a faithful duplication of a prior study and, as such, is considered the “purest” form of replication research in the social sciences. It should be mentioned at this point that a Type I replication is the one most closely associated with the term “replication” in the minds of most researches. More over, this is also the type of replication research most criticized for not being creative. This is somewhat ironic, given the apparent receptivity of reviewers to cross-cultural research that, in many cases, is usually the study of the generalizability of findings from a single country or culture to others and, thus, is simply a Type I replication.

Type II (Similar). A type II replication is a close replication is a close replication of a prior study, and a Type III replication is a deliberate modification of a prior study. Type II replications are the most common form of replication research in marketing settings and are useful in testing phenomena in multiple contexts. If effects are shown in a variety of testing contexts, the case for the findings is strengthened. This has been called the process of triangulation.

Type III (Modification). This replication is a deliberate modification of a prior study. In a Type III replication, the threat of extraneous factors inherent to the nature of human subjects, unless explicitly accounted for in theory testing, is not a factor of concern with regard to replicability.

A.8 Evanschitzky and Armstrong

Where authors define (Evanschitzky y Armstrong, 2010): *Real Replications*. This replication is a duplication of a previously published empirical study that is concerned with assessing whether similar findings can be obtained upon repeating the study. This definition covers what are variously referred to as “exact”, “straight” or “direct” replications. Such works duplicate as closely as possible the research design used in the original study by employing the same variable definitions, settings, measurement instruments, analytical techniques, and so on.

Model Comparisons. This replication is an application of a previously published statistical analysis that is concerned with assessing whether a superior goodness-of-fit can

be obtained, comparing the original statistical model with at least one other statistical model.

Data Re-analyses. This replication can be defined as an application of previously published data that is concerned with assessing whether similar findings can be obtained using a different methodology with the same data or a sub-sample of the data.

A.9 Finifter

Where author defines (Finifter, 1972): *Virtual Replication*. The intention is to repeat an original study not identically but “for all practical purposes” to see whether its results hold up against chance and artifact. Virtual replications are also frequently conducted to find out how dependent a result is on the specific research conditions and procedures used in an original study. To answer this question, one or more of the initial methodological conditions is intentionally altered. For example, a survey or experiment might be repeated except for a change in measuring devices, in the samples used, or in research personnel. If the initial result reappears despite changes, faith in the original finding mounts.

Systematic Replication. The emphasis in systematic replication is not on reproducing either the methods or the substance of a previous study. Instead, the objective is to produce new findings (using whatever methods) which are expected by logical implication to follow from the original study being replicated. When such an implication is actually confirmed by systematic replication, confidence is enhanced not only in the initial finding that prompted the replication but also both in the derived finding and in whatever theoretical superstructure was used to generate the confirmed inference.

Pseudoreplication. It can be defined according to three main operational variations: the repetition of a study on certain subsets of an available total body of real data; the repetition of areal data study on artificial data sets which are intended to simulate the real data; and the repeated generation of completely artificial data sets according to an experimental prescription.

A.10 Finifter

Found in (Finifter, 1975) which it is the same clasification as in (Finifter, 1972).

A.11 Fuess

Fuess (Fuess, 1996) mentions several classifications, basically describing two forms of replication. Replication or duplication aims to run an experiment that resembles the reference experiment as exactly as possible, whereas replication with extension varies the replication to increase external validity (generalizability).

A.12 Hendrick

Where author (Hendrick, 1990) describes the following replications types: *Strict Replication*. An exact, or strict, replication is one in which independent variables (treatments) are duplicated as exactly as possible. That is, the physical procedures are reinstituted as

closely as possible. It is implicitly assumed that contextual variables are either the same as in the original experiment, or are irrelevant.

Partial Replication. A partial replication is some change (deletion or addition) in part of the procedural variables, while other parts are duplicated as in the original experiment. Usually some aspect of the procedures is considered “unessential”, or some small addition is made to expedite data collection.

Conceptual Replication. A conceptual replication is an attempt to convey the same crucial structure of information in the independent variables to subjects, but by a radical transformation of the procedural variables. In addition, specific task variables are often necessarily changed as well.

A.13 Hunter

This author (Hunter, 2001) defines the following replications: *Statistical Replication*. For statistical replications as perfectly replicated studies:

1. All studies measure the independent variable in exactly the same way.
2. All studies measure the dependent variable in exactly the same way.
3. All studies use exactly the same procedure.
4. All studies draw samples from the same population.

Scientific Replication. For scientific replications for simple causal studies:

1. All studies measure the same independent variable X.
2. All studies measure the same dependent variable Y.
3. All studies use essentially the same procedure.
4. All studies should sample from populations that are equivalent in terms of the study question and hence the study outcome. The difference is that statistical replications assume that the word “same” means identical, while scientists interpret the word “same” to mean equivalent.

Conceptual Replication. This replication verifies one of the hypotheses that were not tested in the original study. The researcher of the original study defines control groups to test the most obvious alternative hypotheses against administrative details that are thought to be irrelevant. Any treatment, intervention or manipulation is a set of administrative procedures, which are mostly intrinsic to the active ingredient of the treatment. These replications examine whether the administrative procedures influence the treatments as reflected in the dependent variable.

A.14 Hyman and Wright

Authors (Hyman y Wright, 1967) distinguish three kinds of research operations that are subsumed under the label “replication”.

1. There are studies that build upon, extend, and occasionally test under varying circumstances the findings and hypotheses of previous research.
2. There are studies that attempt to duplicate, as closely as possible, the design, problem, hypotheses, and methods of earlier studies.
3. There is an attempt to conduct a study in several stages, extending it over a relatively long period of time in order to see what new phenomena emerge.

A.15 Van IJzendoorn

Where author (Van IJzendoorn, 1994) defines: *Complete Secondary Analysis*. It is a kind of replication in which all parameters except the researcher and the method of data analysis are kept constant. Secondary analysis also is one of the most inexpensive and efficient types of replication, because it is based on existing data sets. One of the main barriers to secondary replication is, however, the accessibility of the original data sets. The complete secondary analysis may include recoding of the original raw data. In this replication, there are two phases of processing the raw data involved: the coding and analyzing of the data.

Restricted Secondary Analysis. In this type, the coding system is not changed but only the methods of analyzing the data, to see whether the original results survive statistical criticism or the application of refined methods of statistical analysis.

Exact Replication. A replication will be called “exact” if it is essentially similar to the original study. This replication is applied to (dis)confirm the doubts, and to check the assumptions of the varied replications. Many scientists feel that exact replications may be carried out, but usually are irrelevant for scientific progress.

Varied Replication. Replications should be carried out in which new data under different conditions are being collected. From the start, the original study will be “trusted” so much that rather significant variations in the design will be applied. Larger variations may lead to more interesting discoveries in addition to the original study, but they will be followed by smaller variations if more global replications fail to produce new “facts”. If even modest variations fail to reproduce the results, a more or less exact replication is needed.

A.16 Kantowitz et al.

Authors (Kantowitz et al., 1984) describe the following types: *Direct Replication*. This is the attempt to repeat the experiment as closely as is practical, with as few changes as possible in the original method.

Systematic Replication. The experimenter attempts to vary factors believed to be irrelevant to the experimental outcome. If the phenomenon is not illusory, it will survive these changes. If the effect disappears, then the researcher has discovered important boundary conditions on the phenomenon being studied.

Conceptual Replication. One attempts to replicate a phenomenon, but in a way radically different from the original experiment.

A.17 Kelly et al.

Where authors (Kelly et al., 1979) describe: *Literal Replication*. The earlier findings may be reexamined using the same manipulations (independent variables, experimental procedures, etc.) and measures (dependent variables, methods of data analysis, etc.).

Operational Replication. If the experimenter wishes to vary criterion measures, the experiment would be termed an operational replication. In this instance, the dependent variable would represent a different operationalization of the construct; the essential conceptual meaning would remain unchanged.

Instrumental Replication. This replication is carried out when the dependent measures

are replicated and the experimental manipulations are varied. Variations in the implementation of experimental procedures which do not go beyond the originally established relationship would be included in this category.

Constructive Replication. A constructive replication attempt may be identified when both manipulations and measures are varied. This replication involves the attempt to achieve equivalent results using an entirely original methods recipe.

A.18 La Sorte

Author (La Sorte, 1972) distinguishes the following types: *Retest Replication*. In its general form retest replication is a repeat of an original study with few if any significant changes in the research design. The retest has two major purposes: 1) it acts as a reliability check of the original study, and 2) inconsistencies and errors in procedure and analysis can be uncovered in the repeat. Although the retest increases one's confidence that the findings are not artifactual, it does not eliminate the possibility of error in process, especially when the same investigator conducts both studies.

Internal Replication. The differences between the retest and internal replication are mainly procedural. Instead of seeking confirmation of an original study, the internal replication is built into the original study design. So the data, part of which are used for the replication, are gathered simultaneously by the same investigator using a common set of research operations. One finds variations in the procedures for selecting the samples. Two of these procedures are: 1) drawing two or more independent samples, and 2) taking a single sample and later dividing it into subsamples for purposes of analysis and comparison. The internal replication provides an additional data supply which acts to cross-check the reliability of the observed relationships. Thus it is methodologically superior to the single study where the hypothesis is tested only once by one body of data.

Independent Replication. Independent replication is the basic procedure for verifying an empirical generalization. It does this by introducing significant modifications into the original research design in order to answer questions about the empirical generalization that go beyond those of reliability and confirmation. The essential modifications include independent samples drawn from related or different universes by different investigators. These replications differ in design and purpose. They can, however, be broadly categorized into three problem areas. First, is the empirical generalization valid? Second, does further investigation extend it to other social situations or subgroups outside the scope of the original study? Or, third, is the empirical generalization limited by the conditions of particular social situations or specific subgroups?

Theoretical Replication. It involves the inductive process of examining the feasibility of fitting empirical findings into a general theoretical framework. These replications seek to verify theoretical generalizations. In these replications, empirical variables, which have concrete anchoring points are abstracted and conceptualized to a higher theoretical plane, it is necessary to sample a variety of groups using different indicators of the same concepts.

A.19 Leone and Schultz

Where authors (Leone y Schultz, 1980) define: *Experimental Replication*. The same experiment is conducted more than once, although there can be (especially with social systems) no perfect replications. It involves the same method and the same situation.

Nonexperimental Replication. The same method is applied to different situations.

Corroboration. It involves different method and same situation, or different method and different situation.

A.20 Lindsay and Ehrenberg

Where Authors (Lindsay y Ehrenberg, 1993) describe: *Close Replication*. This replication attempts to keep almost all the know conditions of the study much the same or at least very similar (for example, the population or populations in question, the sampling procedure, the measuring techniques, the background conditions, and the methods of analysis). A close replication is particularly suitable early in a program of research to establish quickly and relatively easily and cheaply whether a new result can be repeated at all.

Differentiated Replication. It involves deliberate, or at least known, variations in fairly major aspects of the conditions of the study. The aim is to extend the range of conditions under which the result still holds. Exploring a result with deliberate variations in the conditions of observation is the essence of generalization. According to the authors, there are three reasons for running a differentiated replication:

1. Use different methods (different measuring instruments, analysis procedures, experimental setups, and/or investigators) to reach the same result (triangulation),
2. Extended the scope of the results,
3. Define the conditions under which the generalization no longer holds.

A.21 Lüdtkke

Lüdke (Lüdtkke, 2008) identifies two approaches to replication in the clinical field: *exact replication* and *independent replication*.

Exact replication involves using the same dose of the same medicine for the same length of time and the same type of disease as in the original trial. On the other hand, a replication is *independent* when it is performed independently of the reference study.

Lüdke (Lüdtkke, 2008) gives two examples of clinical trials to illustrate each type of replication. Although the study object of the two trials is the same, they are completely different. The researchers of the two trials have different medical specialities, the research was funded by different sponsors and the studies were run at different hospitals.

A.22 Lykken

This author (Lykken, 1968) defines the following replication types: *Literal Replication*. This involves exact duplication of the first investigator's sampling procedure, experimental conditions, measuring techniques, and methods of analysis.

Operational Replication. One strives to duplicate exactly just the sampling and experimental procedures given in the first author's report. The purpose of operational replication is to test whether the investigator's "experimental recipe" the conditions and procedures he considered salient enough to be listed in the "Methods" section of his report will in other hands produce the results that he obtained.

Constructive Replication. One deliberately avoids imitation of the first author's methods. To obtain an ideal constructive replication, one would provide a competent investigator with nothing more than a clear statement of the empirical "fact" which the first author would claim to have established.

A.23 Mittelstaedt and Zorn

Where authors (Mittelstaedt y Zorn, 1984) describe: *Type I.* The replicating researcher uses the same data sources, models, proxy variables and statistical methods as the original researcher.

Type II. The replicating researcher uses the same data sources, but employs different models, proxy variables and/or statistical methods.

Type III. The replicating researcher uses the same models, proxy variables and statistical methods, but applies them to different data than those used by the original researcher.

Type IV. In this replication, different models, proxy variables and statistical methods are applied to different data.

A.24 Monroe

Where author (Monroe, 1992) defines: *Simultaneous Replication.* Does the same researcher in the same study investigate consumer reactions to more than one product, or to more than one advertisement?

Sequential Replication. Does the researcher or another researcher repeat the study using the same or different stimuli at another point in time?

Nonindependent Replication. The replication is conducted by the same researcher.

Independent Replication. The replication is conducted by different researcher.

Assumed Replication. For example, a researcher using both males and females simultaneously in a study and finding no gender covariate effect assumes replication across gender.

Demonstrated Replication. What is preferable is separate gender conditions wherein the effect has or has not been obtained separately for males and females, that is, demonstrated.

Strict Replication. The replication is a faithful reproduction of the original study.

Partial Replication. The replication is a faithful reproduction of some aspects of the original study.

Conceptual Replication. The replication uses a similar conceptual structure but incorporates changes in procedures and independent variables.

A.25 Radder

This author (Radder, 1992) distinguishes the following types: *Reproducibility of the material realization of an experiment.* In this type of reproduction, the replicator correctly performs all the experimental actions following instructions given by the experimenter who ran the previous experiment. This reproduction is based on a division of labour, where other previously instructed people can run the replication without being acquainted with the theory underlying the experiment. As in this reproduction it is possible to follow the

same procedure to verify the outcome without detailed knowledge of the theory, there may be differences in the theoretical interpretations of the experiment.

Reproducibility of an experiment under a fixed theoretical interpretation. This reproduction implies that the conditions of the previous experiment can be intentionally altered in the replications, provided that the variations are irrelevant to the theoretical interpretation of the experiment.

Reproducibility of the results of an experiment. This type of reproduction refers to when it is possible to achieve the same result as a previous experiment using different methods. This category excludes a reproduction of the same material operationalization.

A.26 Schmidt

Where author (Schmidt, 2009) describes: *Direct Replication*. This involves repeating the procedure of a previous experiment. In this replication, the context variables, the dependent variable or subject selection are open to modification.

Conceptual Replication. This is the use of different methods to retest the hypothesis or result of a previous experiment.

A.27 Sidman

In this classification author (Sidman, 1960) describes the following types:

Direct Replication. This type of replication of a reference experiment is run by the same experimenter. The replication is repeated under the same conditions as the reference experiment and can be run using a between-subjects or within-subjects factors, leading to different single-subject design types.

Systematic Replication. Instead of simply repeating the experiment, data gathered from the original experiment are used as the baseline for running new experiments and thus gathering additional related data. While repeating the same experiment, this replication changes some application condition, that is, the context, one of the subject characteristics, the experimenter, the problem behaviour, etc. A series of successful systematic replications that modify one or more of these factors provides more information about the generality of the outcomes.

A.28 Tsang and Kwan

This authors (Tsang y Kwan, 1999) define the following replication types: *Checking of Analysis*. In this type of replication, the researcher employs exactly the same procedures used in a past study to analyze the latter's data set. Its purpose is to check whether investigators of the original study have committed any errors in the process of analyzing the data.

Reanalysis of Data. Unlike the checking of analysis, in this type of replication, the researcher uses different procedures to reanalyze the data of a previous study. The aim is to assess whether and how the results are affected by problems of definition, as well as by the particular techniques of analysis. Quite often the replication involves using more powerful statistical techniques that were not available when the original study was conducted.

Exact Replication. This is the case where a previous study is repeated on the same population by using basically the same procedures. The objective is to keep the contingent conditions as similar as possible to those of the previous study. The researcher usually uses a different sample of the subjects. The main purpose is to assess whether the findings of a past study are reproducible.

Conceptual Extension. A conceptual extension involves employing procedures different from those of the original study and drawing a sample from the same population. The differences may lie in the way of measuring constructs, structuring the relationships among constructs, analyzing data, and so forth. In spite of these differences, the replication is based on the same theory as the original study. The findings may lead to a revision of the theory.

Empirical Generalization. In this replication, a previous study is repeated on different populations. The researcher runs an empirical generalization to test the extent to which the study results can be generalized to other populations. It follows the original experimental procedures as closely as possible.

Generalization and Extension. The researcher employs different research procedures and draws a sample from a different population of subjects. The more imprecise the replication, the greater the benefit to the external validity of the original finding, if its results support the finding. However, if the result fail to support the original finding, it is difficult to tell whether that lack of support stems from the instability of the finding or from the imprecision of the replication.

Apéndice B

Técnicas y Herramientas Usadas en la Solución

RESUMEN: En este apéndice se describen las técnicas y herramientas empleadas en el procedimiento para comparar replicaciones.

B.1 Mapas Estructurados

Un mapa estructurado (TreeMaps en inglés) (Shneiderman, 1992) es un diagrama que muestra un conjunto de datos de manera jerárquica. Los datos son representados con una serie de rectángulos colocados unos dentro de otros de acuerdo a una estructura de árbol. A cada rama del árbol se le asigna un rectángulo que es dividido en rectángulos más pequeños que representan sub-ramas. A menudo los nodos hoja son coloreados para mostrar una dimensión separada de los datos.

Para representar las particiones con esta técnica, como primer paso se dibuja un rectángulo principal por cada constructo de efecto. A continuación, dentro de cada rectángulo principal se dibujan rectángulos por cada operacionalización de efecto, dentro de cada rectángulo que representa la operacionalización de efecto, se dibujan rectángulos por cada par de constructos de causa. Por último, dentro de cada rectángulo que representa el constructo de causa se dibujan rectángulos por cada par de operacionalizaciones de causa.

Si se desea automatizar esta actividad, se puede usar la herramienta TreeMap versión 4.1.1 (HCIL, 2004) desarrollada por el laboratorio de interacción persona-ordenador HCIL (*Human-Computer Interaction Laboratory*, Laboratorio de Interacción Persona-Ordenador) de la Universidad de Maryland. Ya que las particiones están organizadas de manera jerárquica, esta herramienta facilita la interacción con este tipo de estructura. Por ejemplo, es posible visualizar solamente las particiones que pertenecen a un constructo de efecto. Entre las diversas opciones con las que cuenta esta herramienta, también es posible colorear los diferentes niveles jerárquicos. En la figura B.1 se muestran las particiones coloreadas que pertenecen al constructo referente a la efectividad.

Con la información disponible de la tabla de agrupaciones se genera un fichero de texto delimitado por tabuladores que se suministra a la herramienta para generar las particiones. El fichero de texto utiliza la misma estructura de la tabla de agrupaciones, sin embargo, la herramienta requiere se especifique el tipo de campo de cada columna,

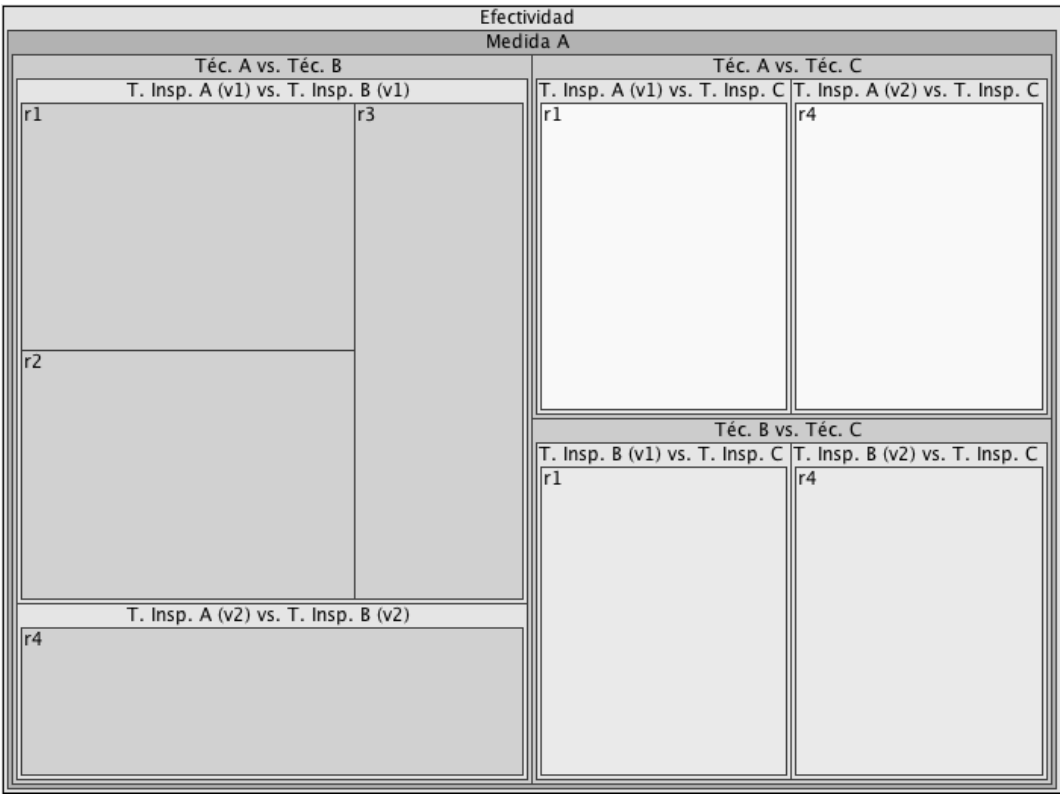
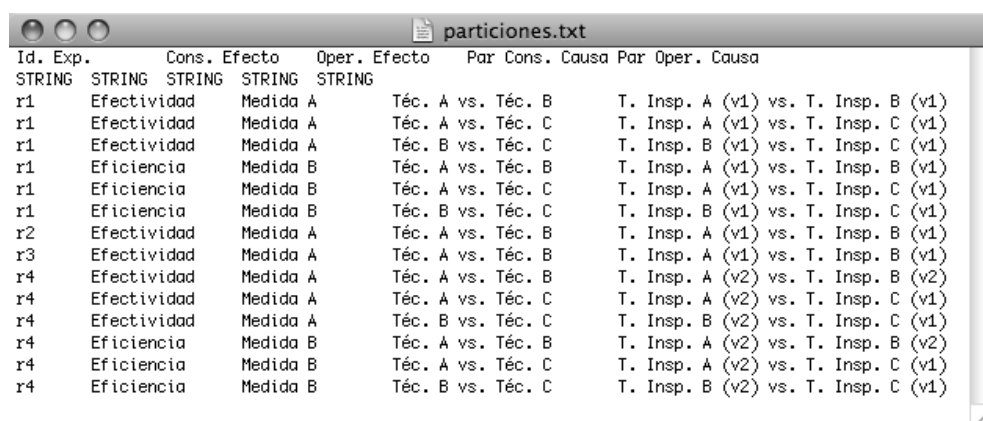


Fig. B.1: Ejemplo de visualización de particiones de acuerdo a la efectividad de varias técnicas

por lo que es necesario añadir una línea para indicar el tipo de campo. Esta línea se añade después de la primera línea que indica los nombres de las columnas. Los campos de las columnas se especifican de tipo **STRING**. en la figura B.2 se muestra la estructura del fichero así como la información de varias particiones.

Al cargar el fichero de texto en la herramienta sólo se visualiza un conjunto de recuadros, donde cada recuadro indica una replicación. En la figura B.3 se muestra la pantalla principal de la herramienta tras cargar el fichero de texto.

Para representar las particiones en la herramienta es necesario especificar el mismo orden jerárquico utilizado en la tabla de agrupaciones. Como primer nivel, se debe especificar los constructos de efecto seguido de las operacionalizaciones de efecto, después se debe especificar el par de constructos de causa y finalmente el par de operacionalizaciones de causa. Para esto, en el panel del costado derecho se debe elegir la opción **Hierarchy** y en el cuadro en el subpanel **Hierarchy** se debe eliminar el elemento por defecto **DEFAULT HIERARCHY**. Una vez eliminado este elemento, en el subpanel **Attributes** se deben añadir en el siguiente orden los nombres de las columnas: *Cons. Efecto, Oper. Efecto, Par Cons. Causa y Par Oper. Causa*. Como resultado se obtiene la visualización mostrada en la figura anterior B.3. Si se desean colorear los recuadros de acuerdo a un nivel determinado se debe seleccionar en el panel principal la opción **Legend**. En esta opción se muestra un subpanel con varias opciones, entre ellas se encuentra la opción para utilizar colores en los recuadros.



Id. Exp.	Cons. Efecto	Oper. Efecto	Par Cons. Causa	Oper. Causa
STRING	STRING	STRING	STRING	STRING
r1	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A (v1) vs. T. Insp. B (v1)
r1	Efectividad	Medida A	Téc. A vs. Téc. C	T. Insp. A (v1) vs. T. Insp. C (v1)
r1	Efectividad	Medida A	Téc. B vs. Téc. C	T. Insp. B (v1) vs. T. Insp. C (v1)
r1	Eficiencia	Medida B	Téc. A vs. Téc. B	T. Insp. A (v1) vs. T. Insp. B (v1)
r1	Eficiencia	Medida B	Téc. A vs. Téc. C	T. Insp. A (v1) vs. T. Insp. C (v1)
r1	Eficiencia	Medida B	Téc. B vs. Téc. C	T. Insp. B (v1) vs. T. Insp. C (v1)
r2	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A (v1) vs. T. Insp. B (v1)
r3	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A (v1) vs. T. Insp. B (v1)
r4	Efectividad	Medida A	Téc. A vs. Téc. B	T. Insp. A (v2) vs. T. Insp. B (v2)
r4	Efectividad	Medida A	Téc. A vs. Téc. C	T. Insp. A (v2) vs. T. Insp. C (v1)
r4	Efectividad	Medida A	Téc. B vs. Téc. C	T. Insp. B (v2) vs. T. Insp. C (v1)
r4	Eficiencia	Medida B	Téc. A vs. Téc. B	T. Insp. A (v2) vs. T. Insp. B (v2)
r4	Eficiencia	Medida B	Téc. A vs. Téc. C	T. Insp. A (v2) vs. T. Insp. C (v1)
r4	Eficiencia	Medida B	Téc. B vs. Téc. C	T. Insp. B (v2) vs. T. Insp. C (v1)

Fig. B.2: Estructura del fichero que contiene la información de las particiones

B.2 Medidas de Proximidad

En la literatura existen diferentes medidas de proximidad que dependiendo del tipo de datos que contengan los objetos a comparar es el tipo de medida a utilizar. Existen medidas para calcular la proximidad de objetos que contienen datos numéricos, medidas para calcular la proximidad de objetos con datos categóricos y datos binarios, y medidas para calcular la proximidad de objetos que contienen una mezcla de varios tipos de datos (Everitt et al., 2001).

Para la identificación de similitudes entre las replicaciones de una partición se usa una medida de proximidad para datos binarios, por lo que los valores de los elementos de la situación experimental de las replicaciones deben convertirse en valores dicotómicos, tal como sucede en la generación de las tablas de particiones. La ventaja de usar datos dicotómicos es que se mantiene un intervalo de valores de 0 a 1 que es intuitivo de comprender en la matriz de similitudes. Estos valores representan una dis-similitud o similitud completa.

Existen varios tipos de medidas de proximidad para datos binarios, algunas de las medidas más conocidas son:

1. Concordancia o acuerdo simple (Sokal y Michener, 1958)
2. Coeficiente de Jaccard (Jaccard, 1908)
3. Rogers and Tanimoto (Rogers y Tanimoto, 1960)
4. Sokal y Sneath (Sokal y Sneath, 1963)
5. Gower y Legendre 1 (Gower y Legendre, 1986)
6. Gower y Legendre 2 (Gower y Legendre, 1986)

Sin embargo, las dos medidas más utilizadas son las dos primeras: acuerdo simple (Sokal y Michener, 1958) y el coeficiente de Jaccard (Jaccard, 1908). Básicamente las medidas 2, 4 y 6 se relacionan de manera monótonica de igual manera sucede con las medidas 1, 3, 5. Es decir, las medidas 4 y 6 son variaciones de la 2 así como las medidas 3 y 5 de la 1.

Las medidas de proximidad para datos binarios se calculan a partir de una tabla de contingencia que contiene los conteos con las igualdades y desigualdades de las variables o contenidos de un par de objetos. Estas medidas se dividen en dos grupos, aquellas que

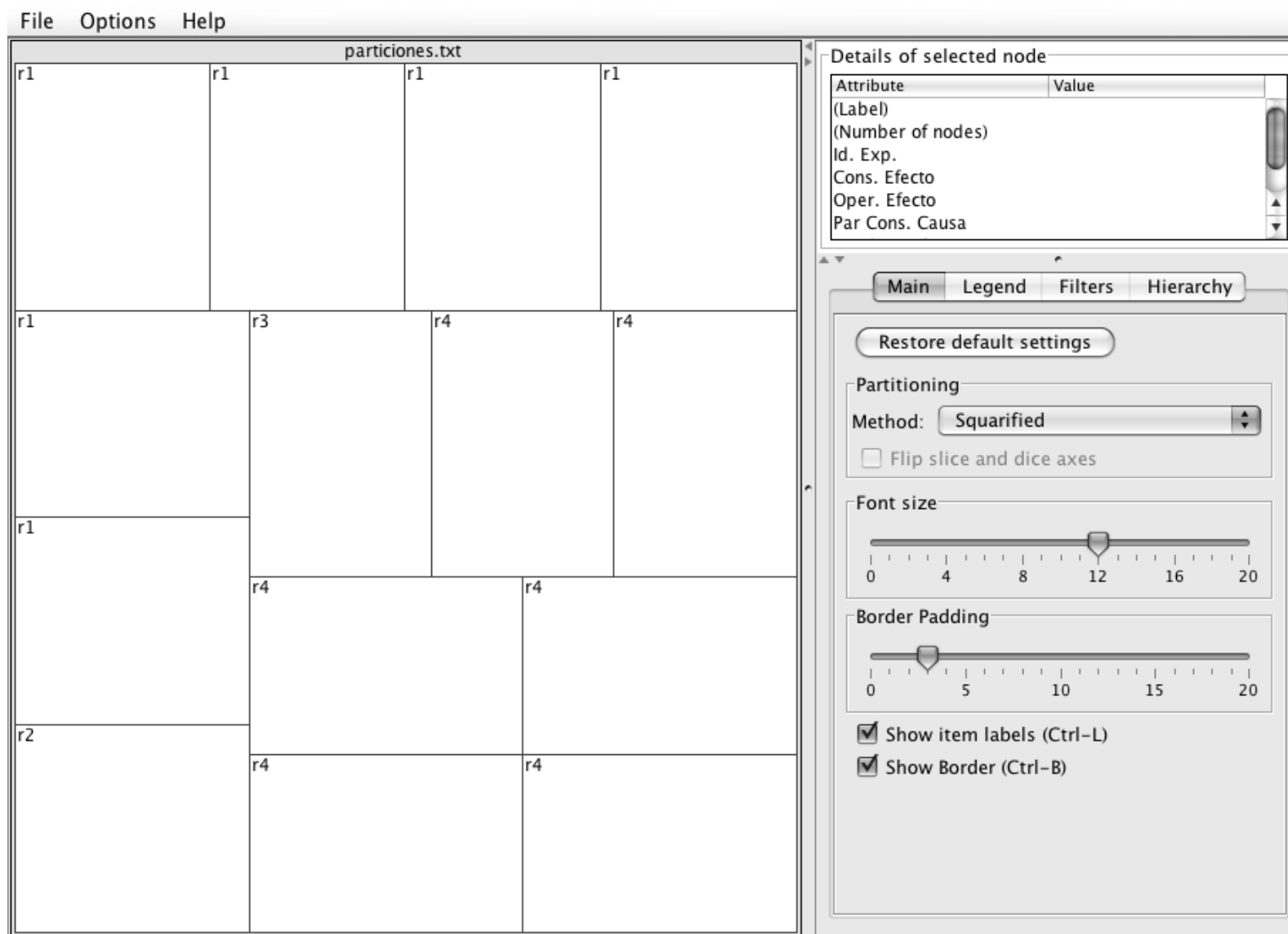


Fig. B.3: Pantalla principal de la herramienta TreeMap

toman en cuenta las desigualdades de las variables (medidas 1, 3 y 5) y las medidas que no toman en cuenta las desigualdades de las variables (medidas 2, 4 y 6). En la tabla B.1 se muestra la estructura de la tabla de contingencia donde se almacenan los conteos de igualdades y desigualdades de las variables de un par de objetos.

objeto _i \ objeto _j	igual (1)	desigual (0)	total
igual (1)	p	q	p+q
desigual (0)	r	s	r+s
total	p+r	q+s	t=p+q+r+s

Tabla B.1: Estructura de la tabla de contingencia donde se almacenan los conteos

En esta tabla de contingencia,

p = es el número de variables dicotómicas que son iguales en ambos Objetos.

q = es el número de variables dicotómicas que son iguales para el Objeto_i y desiguales para el Objeto_j.

r = es el número de variables dicotómicas que son desiguales para el Objeto_i, e iguales para el Objeto_j.

s = es el número de variables dicotómicas que son desiguales para ambos Objetos.

t = número total de variables.

Para clarificar las diferencias entre estos dos grupos de medidas de proximidad, a continuación se describe un ejemplo donde se utilizan las medidas: de acuerdo simple (Sokal y Michener, 1958) y Jaccard (Jaccard, 1908). En la tabla B.2 se muestra la información de dos replicaciones (objeto_i y objeto_j) que usan una determinada técnica de análisis de datos. Cada tipo de técnica representa una variable. Como se observa, ninguna de las replicaciones usa el tipo de análisis ANOVA mixto.

	ANOVA 1 vía	ANOVA 2 vías	ANOVA mixto
experimento A (obj _i)	X		
experimento B (obj _j)		X	

Tabla B.2: Técnicas de análisis de datos usadas en replicaciones A y B

A partir de la tabla de contingencia B.1, se definen las funciones para las medidas de acuerdo simple (Sokal y Michener, 1958) y Jaccard (Jaccard, 1908). Para calcular la medida por acuerdo simple se utiliza la función en (B.1),

$$sim_{ij} = \frac{p + s}{t} \quad (B.1)$$

donde sim_{ij} , representa la similitud resultante entre el objeto_i y objeto_j.

Por otra parte, para calcular el coeficiente de Jaccard (Jaccard, 1908) se utiliza la función en (B.2),

$$sim_{ij} = \frac{p}{p + q + r} \quad (B.2)$$

En la tabla B.3 se muestra una tabla de contingencia con los valores p , q , r y s calculados a partir de la información del par de replicaciones de la tabla B.2. En este caso ninguna replicación comparte un mismo tipo de análisis ($p = 0$) y ambas replicaciones no utilizan el análisis ANOVA mixto ($s = 1$).

objeto, \ objeto,	igual (1)	desigual (0)	total
igual (1)	p=0	q=1	p+q
desigual (0)	r=1	s=1	r+s
total	p+r	q+s	t=p+q+r+s

Tabla B.3: Técnicas de análisis de datos que utilizan dos replicaciones

Empleando la función en (B.1) para calcular la similitud con la medida de acuerdo simple (Sokal y Michener, 1958) y la función en (B.2) para calcular la similitud con la medida Jaccard (Jaccard, 1908) se obtiene para el primer caso,

$$sim_{ij} = \frac{0 + 1}{3} = \frac{1}{3}$$

donde $t = p + q + r + s = 0 + 1 + 1 + 1$

mientras que para el coeficiente de Jaccard (Jaccard, 1908) se obtiene,

$$sim_{ij} = \frac{0}{0 + 1 + 1} = \frac{0}{2}$$

Los valores que se obtienen al utilizar estas funciones se encuentran en el intervalo de 0 a 1, donde 0 representa la ausencia de similitud entre el par de objetos y 1 representa una similitud completa, i.e. los objetos son iguales.

Como se observa en los resultados, al tomar en cuenta las desigualdades comunes en ambas replicaciones (variable s de la tabla de contingencia), la similitud calculada con la medida de acuerdo simple (Sokal y Michener, 1958) aumenta, por otra parte, en el caso del coeficiente de Jaccard (Jaccard, 1908) al no tomar en cuenta las desigualdades comunes (variable s) la similitud resultante es nula.

Para calcular las similitudes entre las replicaciones de una partición se usa el coeficiente de Jaccard (Jaccard, 1908) ya que sólo interesa calcular la similitud de acuerdo a los elementos en común de la situación experimental que hay entre las replicaciones. En este caso, no tiene sentido atribuir un nivel de similitud mayor a un par de replicaciones simplemente porque éstas tengan ausencias en común en los distintos elementos de la situación experimental.

B.3 Cálculo de la Matriz de Similitudes

Una vez descrita la medida de similitud a utilizar, en este caso el coeficiente de Jaccard (Jaccard, 1908), se procede a construir la matriz de similitudes de una partición. Para construir esta matriz se requiere calcular una matriz de similitudes por cada uno de los (sub)elementos de la situación experimental que se registraron en las tablas de particiones. Las matrices resultantes se combinan de forma jerárquica, primero a nivel de

(sub)elemento, después a nivel de secciones hasta obtener una sola matriz de similitudes que contiene las proximidades respecto a los elementos de la situación experimental de las replicaciones de la partición.

La matriz de cada (sub)elemento se construye a partir del cálculo del coeficiente de similitud Jaccard (Jaccard, 1908) de todos los pares de replicaciones, el resultado de cada cálculo se coloca en una matriz simétrica de $N \times N$, donde N representa el número de replicaciones de la partición.

En la tabla B.4 se muestra un ejemplo de tabla de contingencia con el sub-elemento identificado como *técnica de análisis de datos* del elemento *análisis* del protocolo experimental.

	ANOVA 1 vía	ANOVA mixto
r1	X	
r2		X
r3	X	

Tabla B.4: Tabla de contingencia elemento *Técnica de análisis de datos*

A partir de esta tabla se calculan las proximidades con la medida de Jaccard (Jaccard, 1908) descrita en (B.2) obteniendo los siguientes resultados,

$$sim_{er1er1} = \frac{1}{1 + 0 + 0} = \frac{1}{1}$$

$$sim_{er1r1} = \frac{0}{0 + 1 + 1} = \frac{0}{2}$$

$$sim_{er1r2} = \frac{1}{1 + 0 + 0} = \frac{1}{1}$$

$$sim_{r1r1} = \frac{1}{1 + 0 + 0} = \frac{1}{1}$$

$$sim_{r1r2} = \frac{0}{0 + 1 + 1} = \frac{0}{2}$$

$$sim_{r2r2} = \frac{1}{1 + 0 + 0} = \frac{1}{1}$$

Las proximidades resultantes se colocan en una matriz simétrica de $N \times N$. Una característica de las matrices de similitudes es que su diagonal principal se compone de valores numéricos 1. En la tabla B.5 se muestra la matriz de similitudes que representa el sub-elemento *técnica de análisis de datos*.

	r1	r2	r3
r1	1		
r2	0	1	
r3	1	0	1

Tabla B.5: Matriz de similitudes elemento *Técnica de análisis de datos*

Por cada uno de los (sub)elementos y elementos principales de la situación experimental se utiliza el coeficiente de Jaccard (Jaccard, 1908) para calcular el resto de matrices. Por ejemplo, el elemento *diseño* del *protocolo experimental* contiene seis sub-elementos.

Por cada sub-elemento se genera la matriz de similitudes y se calcula la matriz promedio que representa el elemento *diseño*, de igual manera, se promedian las matrices que representan los elementos del *protocolo experimental* para obtener la matriz promedio que representa el protocolo experimental, de manera similar, se promedian las matrices que representan las secciones de la situación experimental para obtener la matriz promedio que representa la situación experimental. En las tablas B.6, B.7, B.8, B.9 y B.10 se muestran las matrices que representan las cinco secciones de la situación experimental: experimentadores, sitio, protocolo experimental, operacionalizaciones y poblaciones.

	r1	r2	r3
r1	1		
r2	0	1	
r3	0	0.7172	1

Tabla B.6: Matriz de similitudes que representa los roles de los experimentadores

	r1	r2	r3
r1	1		
r2	0	1	
r3	0	1	1

Tabla B.7: Matriz de similitudes que representa el sitio

	r1	r2	r3
r1	1		
r2	0.6667	1	
r3	0.9667	0.7	1

Tabla B.8: Matriz de similitudes que representa el protocolo experimental

	r1	r2	r3
r1	1		
r2	1	1	
r3	1	1	1

Tabla B.9: Matriz de similitudes que representa las operacionalizaciones

	r1	r2	r3
r1	1		
r2	1	1	
r3	0.8333	0.8333	1

Tabla B.10: Matriz de similitudes que representa las poblaciones

Como se mencionó, estas matrices se calculan a partir de los (sub)elementos de la situación experimental. Para calcular las matrices antes mostradas, se usaron como ejemplo los valores de los elementos de la situación experimental que a continuación se muestran en las tablas B.11, B.12, B.13 y B.14.

Estructura	Elemento Principal	Categorías	r1	r2	r3
Experimentadores	Diseñador	Experimentador A	X		
		Experimentador C		X	X
	Entrenador	Experimentador A	X		
		Experimentador C		X	X
		Experimentador D			X
	Monitor	Experimentador A	X		
		Experimentador B	X		
		Experimentador C		X	X
		Experimentador D			X
	Medición	Experimentador B	X		
		Experimentador C		X	X
Sitio	Sitio	Experimentador B	X		
		Experimentador C		X	X
		Sitio A	X		
		Sitio B		X	X

Tabla B.11: Ejemplo de elementos en experimentadores y sitio

Estructura	Elemento Principal	Sub Elemento	Categorías	r1	r2	r3
Protocolo experimental	Análisis	Técnica de análisis de datos	ANOVA 1 vía	X		X
			ANOVA mixto		X	
	Diseño	Diseño Experimental	Diseño Aleatorio	X		X
			Diseño Cruzado		X	
		Orden de aplicación de los tratamientos	Aleatorio	X		X
			2 tipos de orden		X	
		Combinación del entrenamiento con las sesiones	Secuencial	X	X	X
		Tratamientos por sesión	3 técnicas, 1 programa	X		
			2 técnicas, 1 programa		X	X
		Sesiones requeridas por sujeto para aplicar tratamiento	1 sesión	X	X	X
		Tareas por sujeto por sesión	1 técnica, 1 programa	X		X
			2 técnicas, 1 programa		X	
	Instrumentos	Instrumentos o equipos de medición	Ordenador, folios, bolígrafo	X	X	X
	Objetos experimentales	Programas (Códigos fuente)	Programa A	X	X	X
		Faltas (instancias)	Instancia de Faltas A	X	X	X
	Materiales	Formularios del experimento	Formularios v1	X	X	X
		Instrucciones del experimento	Instrucciones v1	X	X	X
		Guías suplementarias	Guías v1	X	X	X

Tabla B.12: Ejemplo de elementos en protocolo experimental

Estructura	Elemento Principal	Subelemento	Categorías	r1	r2	r3
Operacionalizaciones de los constructos	(Causas)	Tratamiento A	Tratamiento acorde a la literatura	Inspección dirigida por abstracciones	X	X X
			Procedimiento de aplicación del tratamiento	Proc. de aplicación Téc. A v1	X	X X
			Tratamiento como instrucciones	Instrucciones Téc. A v1	X	X X
			Recursos del tratamiento	Documentación	X	X X
			Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	X	X X
		Tratamiento B	Tratamiento acorde a la literatura	Inspección por casos de uso	X	X X
			Procedimiento de aplicación del tratamiento	Proc. de aplicación Téc. B v1	X	X X
			Tratamiento como instrucciones	Instrucciones Téc. B v1	X	X X
			Recursos del tratamiento	Documentación	X	X X
			Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	X	X X
	(Efectos)	Efectividad	Métrica de la medida faltas observadas	Porcentaje de faltas observadas	X	X X
			Procedimiento de Medición	Proc. A	X	X X

Tabla B.13: Ejemplo de elementos en operacionalizaciones de los constructos

Estructura	Elemento	Subelemento	Categorías	r1	r2	r3
Poblaciones	Propiedades de los sujetos	Tipo de sujeto	Pregrado	X	X	
			Grado (Máster)			X
		Experiencia en técnica	Solo para experimento	X	X X	
		Otras condiciones del sujeto (conocimiento del lenguaje)	Apto para el experimento	X	X X	
	Propiedades de los objetos	Lenguaje	Java	X	X X	
		Tamaño	300LOC	X	X X	
		Esquema de faltas	Esquema A	X	X X	

Tabla B.14: Ejemplo de elementos en poblaciones

Al promediar las distintas matrices relacionadas con los cinco elementos principales de la situación experimental, se obtiene la matriz de similitudes que representa la situación experimental completa, esta matriz se muestra en la tabla B.15.

	r1	r2	r3
r1	1		
r2	0.5333	1	
r3	0.56	0.8501	1

Tabla B.15: Matriz de similitudes que representa la situación experimental de una partición

Una vez que se dispone de la matriz de similitudes que representa la situación experimental, es necesario transformar esta matriz en una matriz de distancias¹ cuyas distancias estén representadas en un espacio euclídeo, esto con el fin de tener una interpretación de las similitudes como distancias físicas (Gower y Legendre, 1986). Existen varias funciones para convertir la matriz de similitudes en una matriz de distancias en un espacio euclídeo (Everitt et al., 2001), cualesquiera de las siguientes funciones (B.3), (B.4) y (B.5) satisfacen este propósito,

$$d_{ij} = \sqrt{1 - s_{ij}} \quad (\text{B.3})$$

donde d representa la matriz de distancias resultante y s la matriz de similitudes de la situación experimental,

$$d_{ij} = \sqrt{\delta_{ij}^2 + c_1} \quad (\text{B.4})$$

donde d representa la matriz de distancias resultante, δ la matriz de dis-similitudes $(1 - s)$ y c_1 constante de ajuste (Lingoes, 1971),

$$d_{ij} = \delta_{ij} + c_2 \quad (\text{B.5})$$

donde d representa la matriz de distancias resultante, δ la matriz de dis-similitudes $(1 - s)$ y c_2 constante de ajuste (Cailliez, 1983).

En la tabla B.16 se muestra la matriz de distancias en un espacio euclídeo tras aplicar la función en (B.3).

¹Este tipo de matriz también se le conoce como matriz de dis-similitud.

	r1	r2	r3
r1	0		
r2	0.6831	0	
r3	0.6633	0.3872	0

Tabla B.16: Matriz de distancias en un espacio euclídeo que representa la situación experimental

B.4 Análisis de Conglomerados Jerárquico

El análisis de conglomerados jerárquico es una técnica estadística que se usa para clasificar objetos en grupos relativamente homogéneos llamados conglomerados. Los objetos en cada grupo tienden a ser similares entre sí y diferentes a los objetos en otros grupos. En este tipo de análisis se utiliza un árbol llamado *dendrograma* como representación gráfica para interpretar los resultados.

En la literatura existen varios algoritmos para formar los conglomerados (Everitt et al., 2001). Para la identificación de similitudes se usa el algoritmo jerárquico aglomerativo. Este algoritmo produce una sucesión de conglomerados de objetos de tal manera que en cada iteración el número de conglomerados va disminuyendo hasta que se forma un sólo conglomerado. Este algoritmo puede verse como un enfoque ascendente ya que se parte de varios conglomerados donde éstos se van agrupando en otros conglomerados hasta formar uno sólo.

Existen varias funciones para calcular las distancias que hay entre los conglomerados al momento de unir o formar nuevos conglomerados. Los tres tipos de distancias que se pueden calcular entre un par de conglomerados A y B son:

- Distancia mínima entre los elementos de cada conglomerado (Sneath, 1957) (a esta función se le conoce también como enlace sencillo o vecino más cercano)

$$\min\{d(x, y) : x \in A, y \in B\}$$

- Distancia máxima entre los elementos de cada conglomerado (Sørensen, 1948) (a esta función se le conoce también como enlace completo o vecino más lejano)

$$\max\{d(x, y) : x \in A, y \in B\}$$

- Distancia promedio entre los elementos de cada conglomerado (Sokal y Michener, 1958) (a esta función se le conoce también como promedio de grupo)

$$\frac{1}{|A| \cdot |B|} \sum_{x \in A} \sum_{y \in B} d(x, y)$$

La función de distancia que se usa para esta actividad del sub-proceso *identificación de similitudes* es la distancia promedio (Sokal y Michener, 1958) ya que esta función tiende a unir los conglomerados con varianzas pequeñas. Es una función intermedia entre la distancia máxima y la mínima por lo que es relativamente robusta (Everitt et al., 2001). Para formar los conglomerados y obtener el dendrograma se siguen los pasos que a continuación se describen. Como ejemplo en la tabla B.17 se muestra la matriz de distancias respecto a los elementos de la situación experimental de tres replicaciones *r1*, *r2* y *r3* pertenecientes a una partición.

	r1	r2	r3
r1	0		
r2	0.6831	0	
r3	0.6633	0.3872	0

Tabla B.17: Matriz de distancias de las replicaciones $r1$, $r2$ y $r3$

1. Como primer paso, se señalan en la matriz los elementos con la distancia mínima, en este ejemplo son $r2$ y $r3$ que tienen una distancia de 0.3872.
2. Estos elementos se reemplazan en la matriz por el conjunto que forman. En este caso, $[r2, r3]$.
3. La distancia de este conjunto con un elemento de la matriz se calcula como, el promedio de las distancias entre ese elemento y cada uno de los elementos del conjunto. Aquí, por ejemplo la distancia entre $[r2, r3]$ y $r1$ es: $\text{promedio}((r2, r1), (r3, r1)) = \text{promedio}(0.6831, 0.6633) = 0.6732$.

Al finalizar este cálculo con los diferentes elementos de la matriz se obtiene una nueva matriz como se muestra en la tabla B.18.

	[r2,r3]	r1
[r2,r3]	0	
r1	0.6732	0

Tabla B.18: Matriz resultante tras la primera iteración de formación de conglomerados

El algoritmo termina cuando ya no es posible generar más conglomerados, como sucede en este ejemplo. En caso de haber más elementos en la matriz, se realiza una nueva iteración comenzando en el paso 1 hasta obtener una matriz como la que se muestra en la tabla B.18. Las matrices que se generan en cada iteración van almacenando las distancias que hay entre los conglomerados, con esta información se construye el dendrograma. En este ejemplo la distancia que hay entre el primer conglomerado $[r2, r3]$ es de 0.3872 unidades y la distancia que hay entre este conglomerado y $r1$ es de 0.6732 unidades. En la figura B.4 se muestra el dendrograma que se obtiene con la información de las matrices generadas en cada iteración. En el dendrograma se muestran los diferentes conglomerados generados por este algoritmo, en este caso 2.

B.5 Análisis de Escalado Multidimensional

Como otra alternativa para examinar las similitudes entre las replicaciones de una partición, también se realiza el análisis de escalado multidimensional como parte de la identificación de similitudes. Al igual que el análisis de conglomerados jerárquico, el escalado multidimensional o MDS es una técnica estadística multivariante que permite apreciar de forma gráfica el grado de semejanza de un conjunto de objetos, en este caso, replicaciones.

El MDS es una técnica de representación espacial que trata de visualizar sobre un mapa un conjunto de objetos cuya posición relativa se desea analizar. El propósito del MDS es transformar una matriz de proximidad en distancias susceptibles de ser representadas en un espacio multidimensional. El MDS está basado en la comparación de objetos, de forma que si se juzga a los objetos A y B como los más similares entonces esta técnica colocará a los objetos A y B en el gráfico de forma que la distancia entre ellos sea más pequeña que la distancia entre cualquier otro par de objetos.

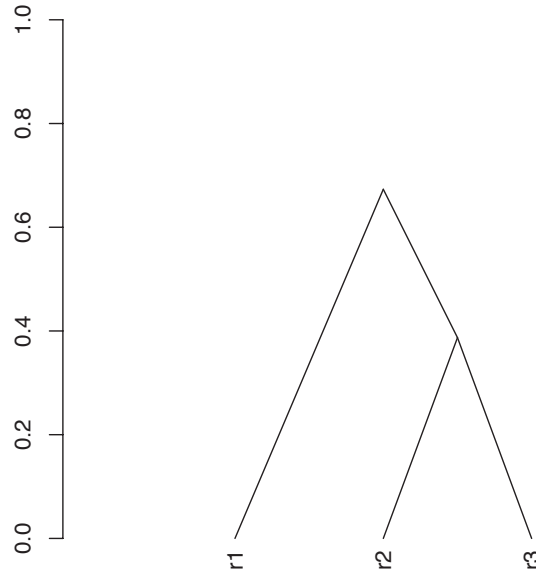


Fig. B.4: Dendrograma resultante tras aplicar el algoritmo de conglomerados jerárquico

En esta actividad se utiliza el escalado multidimensional métrico también conocido como análisis de coordenadas principales o escalado clásico. En esta técnica se parte de una matriz de distancias Δ y se busca una configuración Y , en un espacio euclídeo (\mathbb{R}^q) en el que las distancias entre los puntos sean las contenidas en Δ . Básicamente esta técnica consiste en transformar una matriz de distancias Δ , en una matriz de productos escalares B , y transformar la matriz de productos escalares en una matriz de coordenadas Z .

En esta técnica se suelen hacer representaciones gráficas de dos y tres dimensiones para comprender la configuración resultante i.e. la matriz de coordenadas Z . A continuación se describen los pasos para representar las similitudes entre las replicas de una partición en un mapa de dos dimensiones a través de esta técnica².

1. A partir de la matriz de distancias $\Delta = (\delta_{ij})$ (ver función B.3)³ que representa los elementos de la situación experimental de las replicas de una partición, en este caso $r1$, $r2$ y $r3$ construir la matriz A con la función en (B.6),

$$\delta_{ij} = \begin{pmatrix} 0 & 0.6831 & 0.6633 \\ 0.6831 & 0 & 0.3872 \\ 0.6633 & 0.3872 & 0 \end{pmatrix}$$

$$A = \left(-\frac{1}{2} \delta_{ij}^2 \right) \quad (\text{B.6})$$

²Para conocer en profundidad el funcionamiento de esta técnica, consultar el artículo de Gower (Gower, 1966).

³Ya que en esta técnica se realizan varios cálculos matriciales es necesario que la matriz de distancias contenga todos sus valores i.e. debe tener los valores que componen el triangulo superior de la matriz.

$$A = \begin{pmatrix} 0 & -0.23331281 & -0.21998344 \\ -0.2333128 & 0 & -0.07496192 \\ -0.2199834 & -0.07496192 & 0 \end{pmatrix}$$

2. Obtener la matriz B cuyos elementos son $b_{ij} = a_{ij} - \bar{a}_{i\bullet} - \bar{a}_{\bullet j} + \bar{a}_{\bullet\bullet}$ i.e. restando a cada elemento de A la media de su fila y la de su columna y sumando la media de todos los elementos. Tras calcular B se obtiene la matriz,

$$B = \begin{pmatrix} 0.18480678 & -0.096846516 & -0.087960269 \\ -0.09684653 & 0.088125784 & 0.008720741 \\ -0.08796026 & 0.008720731 & 0.079239528 \end{pmatrix}$$

3. Realizar la descomposición espectral de B , $B = V\Lambda V'$ donde Λ es la matriz diagonal con los valores propios de B , y V es una matriz ortogonal $n \times r$ con los r vectores propios asociados a los valores propios de B en sus columnas. Realizando estos cálculos se obtiene la matriz de valores propios y la matriz de vectores propios,

$$\Lambda = \begin{pmatrix} 0.277502 & 0 & 0 \\ 0 & 0.074669 & 0 \\ 0 & 0 & -6.666 \times 10^{-10} \end{pmatrix}$$

$$V = \begin{pmatrix} 0.8159078 & -0.03100115 & -0.5773503 \\ -0.4348018 & -0.69109633 & -0.5773503 \\ -0.3811061 & 0.72209749 & -0.5773503 \end{pmatrix}$$

4. La matriz con las coordenadas principales es entonces $Z = V\Lambda^{1/2}$, donde las coordenadas buscadas se obtienen de las primeras columnas de los vectores propios V , en este caso se seleccionan los dos primeros vectores propios para calcular las coordenadas en un espacio de dos dimensiones. Usando la función en (B.7) se obtiene,

$$Z = V\Lambda^{1/2} \quad (\text{B.7})$$

$$Z = \begin{pmatrix} 0.8159078 & -0.03100115 \\ -0.4348018 & -0.69109633 \\ -0.3811061 & 0.72209749 \end{pmatrix} \begin{pmatrix} \sqrt{0.277502} & 0 \\ 0 & \sqrt{0.074669} \end{pmatrix}$$

$$= \begin{pmatrix} 0.4298081 & -0.008471288 \\ -0.2290471 & -0.188847064 \\ -0.2007610 & 0.197318355 \end{pmatrix}$$

De acuerdo a los resultados de Z en la tabla B.19 se muestra la tabla con las coordenadas de las replicasiones $r1$, $r2$ y $r3$.

5. Este último paso consiste en realizar el gráfico en dos dimensiones de las coordenadas resultantes, en la figura B.5 se muestra la gráfica con las coordenadas de las replicasiones, como se observa $r1$ se encuentra más alejado a $r2$ y $r3$, por lo que se interpreta que $r1$ es menos similar a $r2$ y $r3$.

experimento	x	y
r1	0.4298081	-0.0084713
r2	-0.229047	-0.1888471
r3	-0.200761	0.1973184

Tabla B.19: Coordenadas resultantes de cada replicación

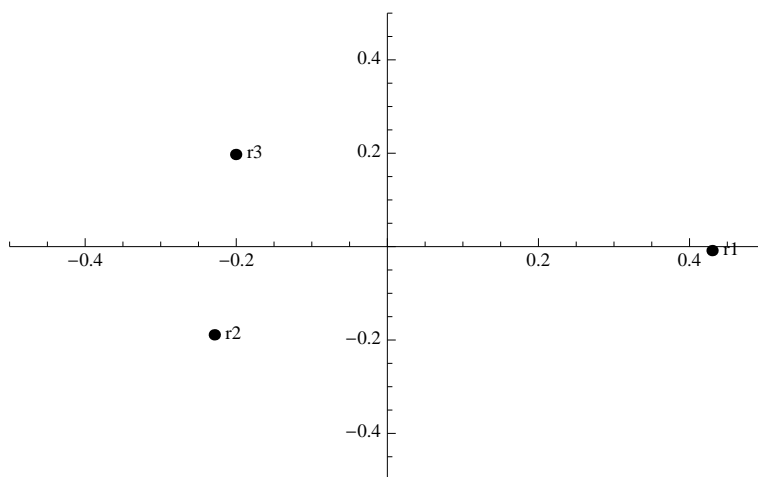


Fig. B.5: Gráfico en dos dimensiones con las coordenadas resultantes del MDS métrico

B.6 Automatización de Cálculos para la Identificación de Similitudes

Con el fin de reducir el esfuerzo y tiempo que conlleva calcular de forma manual las matrices relacionadas con los diferentes (sub)elementos de la situación experimental, combinar las matrices resultantes para obtener la matriz que representa la situación experimental y realizar los análisis de conglomerados jerárquico así como el análisis de escalado multidimensional, se ha automatizado la elaboración de estos cálculos.

Se ha escrito un programa que genera de forma automática las tablas de contingencia de cada uno de los (sub)elementos de la situación experimental registrados en la tabla de partición. Por otra parte, para el cálculo de las matrices de distancias y la combinación de éstas se ha utilizado el entorno y lenguaje de programación R (R Dev. Core Team, 2011). A través de un fichero de texto que contiene varias funciones escritas en este lenguaje se ejecuta para obtener la matriz de distancias que representa la situación experimental de las replicaciones de una partición. Este mismo fichero se utiliza para realizar los análisis de conglomerados jerárquico y de escalado multidimensional. A continuación se describen los pasos para realizar esta automatización.

B.6.1 Generación de Tablas de Contingencia

Como primer paso es necesario agrupar la información de las diferentes tablas que conforman la tabla de partición (formulario PCR-6 en el apéndice D.6) en un sola tabla. Esta tabla que se almacena en un fichero de texto se ingresa como parámetro al programa encargado de generar las tablas de contingencia. Este programa genera una tabla de contingencia por cada (sub)elemento de la situación experimental.

Para construir esta tabla que representa la tabla de partición, se copia el identificador

	A	B	C	D		A	B	C	D
1	ID	r1	r2	r3	28	3.2.6		X	
2	1.1	X			29	3.3.1	X	X	X
3	1.1		X	X	30	3.4.1	X	X	X
4	1.2	X			31	3.4.2	X	X	X
5	1.2		X	X	32	3.5.1	X	X	X
6	1.2			X	33	3.5.2	X	X	X
7	1.3	X			34	3.5.3	X	X	X
8	1.3	X			35	4.1.1.1	X	X	X
9	1.3		X	X	36	4.1.1.2	X	X	X
10	1.3			X	37	4.1.1.3	X	X	X
11	1.4	X			38	4.1.1.4	X	X	X
12	1.4		X	X	39	4.1.1.5	X	X	X
13	1.5	X			40	4.1.2.1	X	X	X
14	1.5		X	X	41	4.1.2.2	X	X	X
15	2	X			42	4.1.2.3	X	X	X
16	2		X	X	43	4.1.2.4	X	X	X
17	3.1.1	X		X	44	4.1.2.5	X	X	X
18	3.1.1		X		45	4.2.1	X	X	X
19	3.2.1	X		X	46	4.2.2	X	X	X
20	3.2.1		X		47	5.1.1	X	X	
21	3.2.2	X		X	48	5.1.1			X
22	3.2.2		X		49	5.1.2	X	X	X
23	3.2.3	X	X	X	50	5.1.3	X	X	X
24	3.2.4	X			51	5.2.1	X	X	X
25	3.2.4		X	X	52	5.2.2	X	X	X
26	3.2.5	X	X	X	53	5.2.3	X	X	X
27	3.2.6	X		X	54				

Fig. B.6: Estructura del fichero con la información de la tabla de partición

de cada (sub)elemento así como las replicaciones de la partición con la que se está trabajando en una hoja de cálculo. En la figura B.6 se muestra un ejemplo de la estructura de una tabla de partición con la información de los elementos de la situación experimental de las replicaciones que conforman una partición. Los identificadores numéricos indican el (sub)elemento de la situación experimental. En este fichero, el primer encabezado de la tabla debe comenzar con el identificador ID escrito en mayúsculas seguido de los identificadores de las replicaciones.

Una vez que se tiene la tabla completa, se guarda el fichero como valores separados por comas especificando la extensión del fichero como CSV. Por ejemplo, el fichero se ha guardado con el nombre `sitexp_part01.csv`.

El siguiente paso consiste en generar una tabla de contingencia por cada (sub)elemento de la situación experimental. Para esto se ejecuta el programa denominado **Transpose** cuyo código fuente se encuentra en el apéndice C.1 de este documento. El programa **Transpose** recibe como entrada un fichero de texto CSV que contiene la información de la tabla de partición y genera varias tablas de contingencia que almacena en ficheros de texto. Cada fichero contiene una tabla de contingencia de uno de los (sub)elementos de la situación experimental que se encuentran en la tabla de partición del fichero CSV.

Este programa está escrito en el lenguaje de programación Java por lo que es necesario

contar con el entorno de ejecución Java ⁴. El programa **Transpose** utiliza un paquete que contiene utilerías para el manejo de ficheros de texto (Infomata, 2009) por lo que es necesario leer un fichero JAR (*Java Archive*) que contiene estas utilerías al momento de su ejecución. El programa se debe ejecutar en modo línea de comandos del sistema operativo.

A continuación se describen algunas formas de ejecutar el programa tanto en sistemas operativos *nix ⁵ como en Windows. Se asume que el fichero **Transpose.class**, **datafile.jar** y en este caso el fichero **sitexp_part.csv** que contiene la tabla de partición, se encuentran dentro de un mismo directorio.

1. `java -cp '..datafile.jar' Transpose sitexp_part01.csv`
2. `java -cp C:\part01\contingency_tables\datafile.jar;
C:\part01\contingency_tables Transpose sitexp_part01.csv`
3. `java Transpose sitexp_part01.csv`

La primera forma funciona en los sistemas operativos *nix, mientras que la segunda forma funciona en el sistema operativo Windows ⁶. La tercera forma funciona en ambos sistemas operativos siempre y cuando la variable de entorno **CLASSPATH** apunte al directorio donde se encuentran los ficheros **.jar** y **.class**. En la figura B.7 se muestran los ficheros generados tras ejecutar el programa **Transpose**.

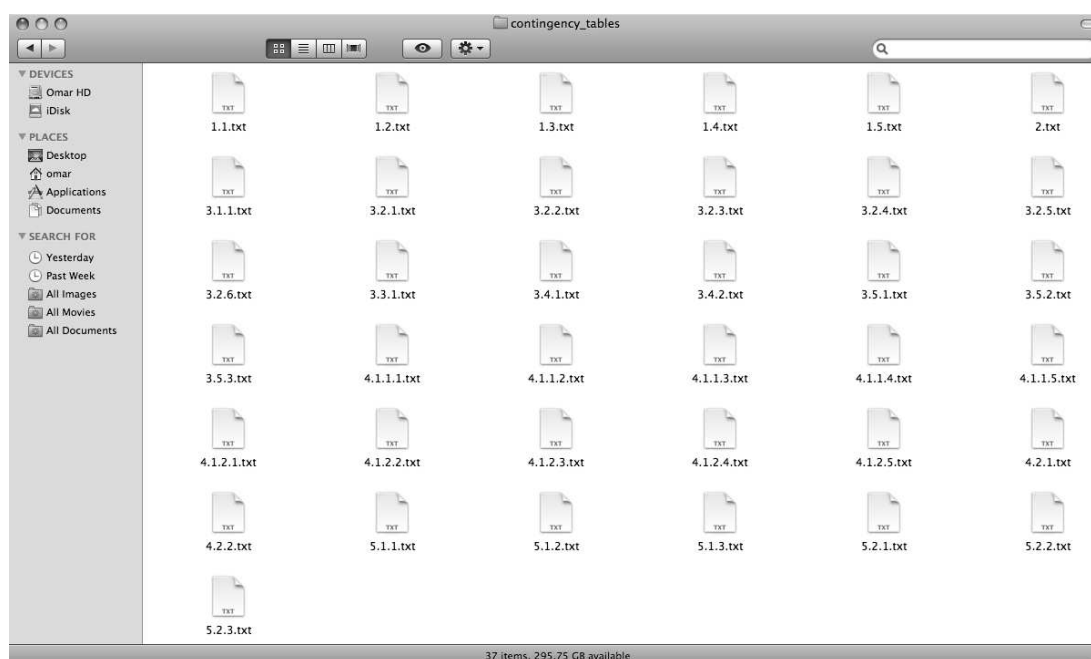


Fig. B.7: Ficheros de texto generados por el programa **Transpose**

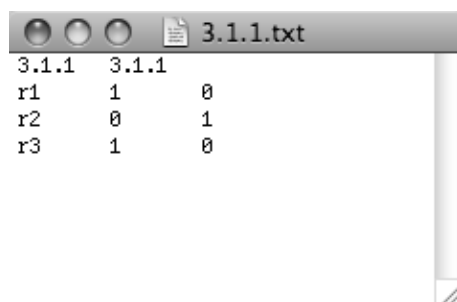
Los ficheros de texto generados se nombran de acuerdo al identificador de cada (sub)elemento de la situación experimental. Cada fichero contiene la tabla de contingencia que se utilizará para calcular las diferentes matrices de distancia. En la figura B.8 se muestra el contenido de uno de los ficheros generados por el programa **Transpose**, en

⁴El programa **Transpose** está compilado con el JDK (*Java Development Kit*) versión 1.5.0_26.

⁵Este término se refiere a todos los sistemas operativos del tipo Unix.

⁶Se asume que los ficheros se encuentran en el directorio **C:\Transpose**.

este caso el fichero relacionado con el sub-elemento Técnica de análisis de datos con ID 3.1.1.



3.1.1	3.1.1	
r1	1	0
r2	0	1
r3	1	0

Fig. B.8: Fichero de texto con la tabla de contingencia del (sub)elemento 3.1.1

Como se observa en la figura B.8, lo que hace el programa es transponer la información y sustituir el símbolo “X” por el valor numérico 1 y sustituir las ausencias de valor por el valor numérico 0.

B.6.2 Cálculo de matriz de distancias

Una vez que se tienen las tablas de contingencia contenidas en los diferentes ficheros de texto, el siguiente paso consiste en calcular una matriz de similitudes por cada tabla de contingencia y combinar las matrices de forma jerárquica: a nivel de sub-elementos, a nivel de elementos, y a nivel de secciones para obtener una sola matriz que contiene las similitudes de las replicaciones de la partición respecto a los elementos de la situación experimental. Una vez que se dispone de esta matriz se convierte en una matriz de distancias en un espacio euclídeo que se utiliza para realizar los análisis de conglomerados jerárquico y de escalado multidimensional.

Esta serie de cálculos se realiza en el entorno R (R Dev. Core Team, 2011) ejecutando un fichero de texto que contiene varias funciones escritas en este lenguaje de programación. R es un lenguaje y entorno de programación multiplataforma para el análisis estadístico y gráfico. Este lenguaje forma parte de un proyecto colaborativo y abierto donde sus usuarios pueden publicar paquetes que extienden su configuración básica. Los paquetes contienen diferentes funciones para realizar diversos cálculos o gráficos. Para el cálculo de las matrices se emplea el paquete *ade4* (Dray y Dufour, 2007; Chessel et al., 2004; Dray et al., 2007) que entre otras funciones contiene una función para calcular la medida de similitud Jaccard (Jaccard, 1908) en datos binarios, y funciones para convertir la matriz de similitudes resultante en una matriz de distancias en un espacio euclídeo. Para calcular las matrices de distancia se ejecuta el fichero de texto denominado **r_script01.txt** que contiene varias funciones para realizar estos cálculos, el contenido de este fichero se encuentra disponible en el apéndice C.2 de este documento.

La carga del fichero se realiza a través de la función **source()** que como parámetro recibe la ubicación y el nombre de un fichero de texto⁷.

Continuando con el ejemplo, en la figura B.9 se muestra el entorno R y la carga del fichero de texto que se encuentra dentro de un directorio llamado **part01/sim_comp**⁸.

⁷En este caso, **part01/sim_comp/r_script01.txt**.

⁸En este ejemplo el espacio de trabajo del entorno R se ha configurado para que acceda al contenido de este directorio.

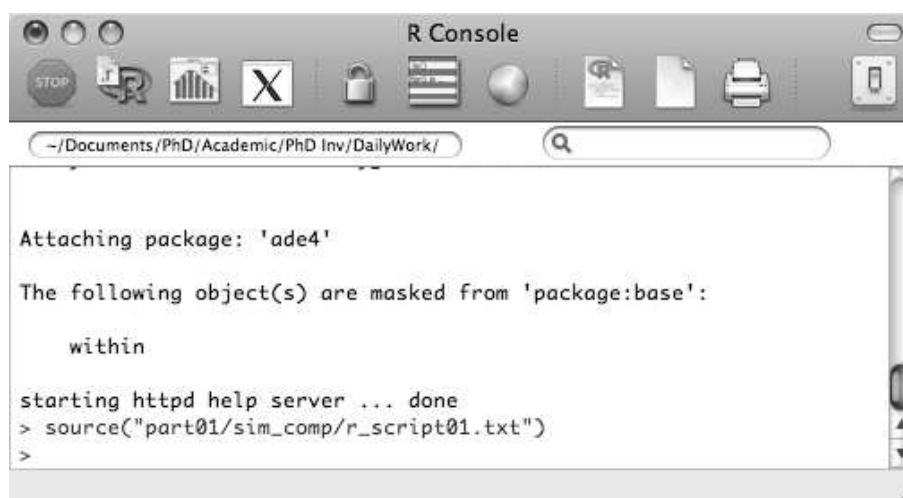


Fig. B.9: Carga del fichero que contiene las funciones para calcular las matrices de distancias

Entre las distintas variables que se utilizan en las funciones del fichero de texto, se encuentra la variable `ec` que contiene la matriz resultante de las combinaciones del resto de matrices. En la figura B.10 se muestra el resultado de esta matriz de distancias. Esta matriz se utiliza posteriormente para realizar los análisis de conglomerados jerárquico y escalado multidimensional.

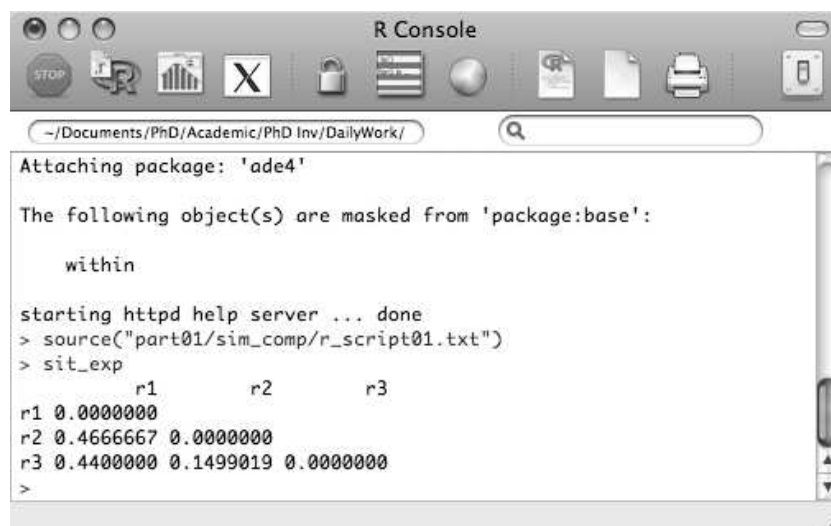


Fig. B.10: Matriz de distancias de la situación experimental de las replicaciones *er1*, *r1* y *r2*

Como se observa en la figura B.10, los resultados de esta matriz indican que la replicación *r2* y *r3* tienen una mayor similitud respecto a la replicación *r1*.

B.6.3 Conglomerados Jerárquico y Escalado Multidimensional

El fichero de texto que se utiliza para calcular las diferentes matrices de distancias, contiene también las funciones para realizar estos dos análisis, por lo que cuando se ejecuta este fichero en el entorno R (R Dev. Core Team, 2011) se realizan los análisis de conglomerados jerárquico y de escalado multidimensional métrico obteniendo como resultado dos ficheros que contienen las representaciones visuales de ambos análisis.

Al ejecutar el fichero de texto con la función `source()`⁹ se realizan los cálculos de las matrices de distancias descritos anteriormente así como los cálculos relacionados con los análisis de conglomerados jerárquico y escalado multidimensional. Como resultado, se generan dos ficheros EPS¹⁰ que contienen las representaciones visuales de las replicaciones $r1$, $r2$ y $r3$ que pertenecen a la partición seleccionada. En la figura B.11 se muestran las representaciones gráficas resultantes. La figura B.11b contiene una cuadrícula donde el valor d ayuda a interpretar las distancias (similitudes) entre las replicaciones de la partición.

B.7 Gráficos de Control

Inventado por Shewhart (Shewhart, 1931), los gráficos de control son una herramienta comúnmente usada en empresas del sector industrial. Básicamente un gráfico de control representa los valores de algún tipo de medición realizada durante el funcionamiento de un proceso continuo y que sirve para controlar dicho proceso. En un funcionamiento normal del proceso, los valores fluctúan dentro de un intervalo determinado, en caso de haber alguna anomalía en el proceso los valores salen de este intervalo.

Para generar el gráfico de control es necesario estimar el promedio así como los límites bajo los cuales las observaciones se consideran como fluctuaciones comunes del proceso. Se asume que las observaciones se distribuyen de acuerdo a una distribución normal, por lo que se dirá que el proceso está bajo control si las observaciones obtenidas siguen las características de una distribución normal. Aproximadamente el 67% de las observaciones han de encontrarse dentro de la primera desviación ($\pm 1\sigma$), aproximadamente el 95% de las observaciones han de encontrarse dentro de la segunda desviación ($\pm 2\sigma$) y aproximadamente el 99% de las observaciones han de encontrarse dentro de la tercera desviación ($\pm 3\sigma$). En esta última desviación se delimita el proceso con los límites superior UCL e inferior LCL.

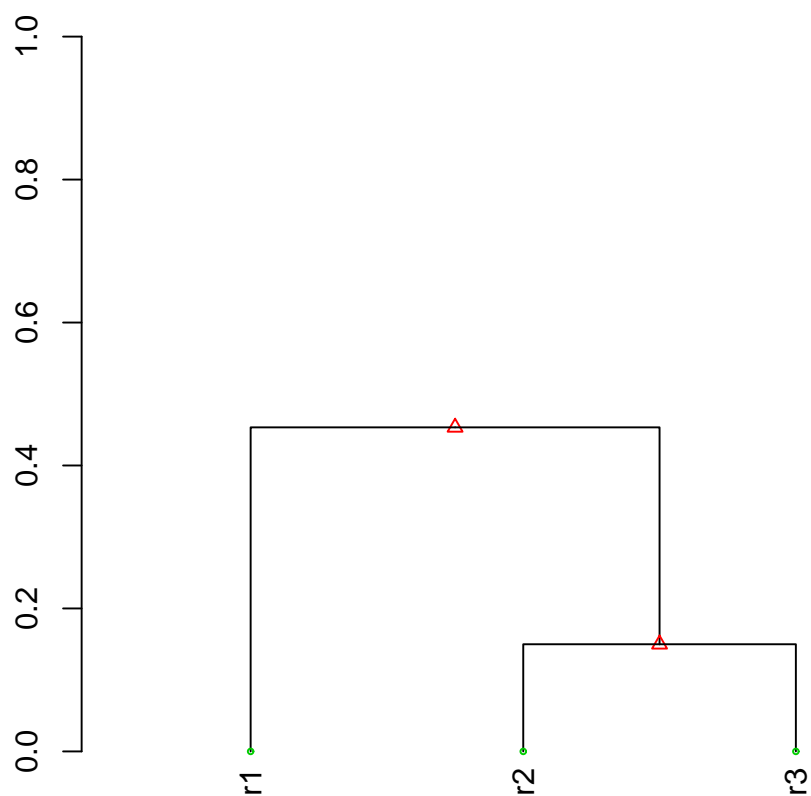
Un gráfico de control contiene tres zonas que ayudan a identificar si un proceso está bajo control o fuera de control (Oakland, 2003).

- En la primera zona (zona 1, zona C o zona estable) las observaciones se encuentran estables, las diferencias de magnitud observadas se deben a la variabilidad común del proceso¹¹.
- En la segunda zona (zona 2, zona B o zona de alerta) se debe ser cuidadoso, ya que el proceso puede estar mostrando causas especiales de variación.
- En la tercera zona o más allá de ésta (zona 3, zona A o zona de acción) se debe

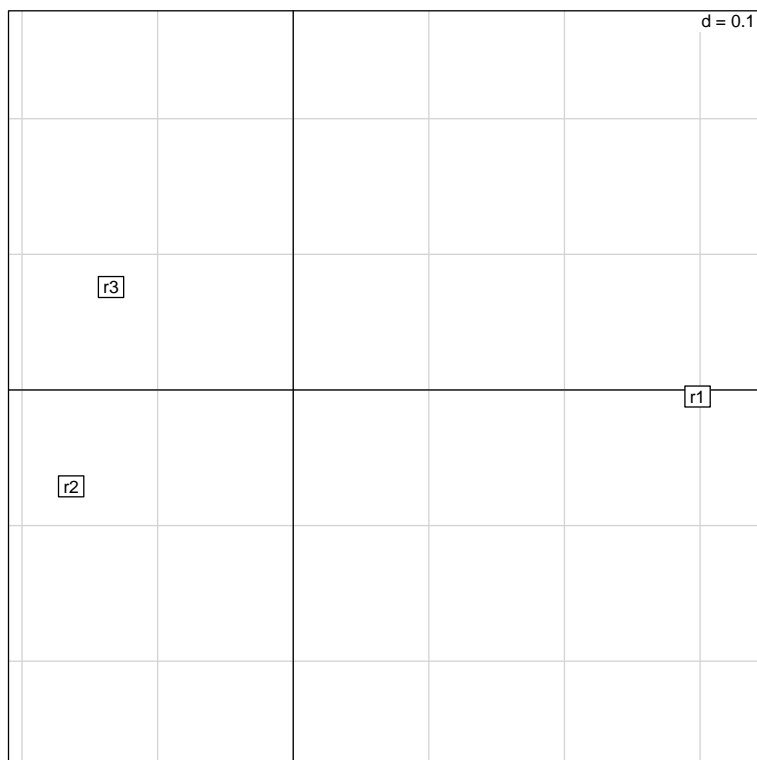
⁹`source("part01/sim_comp/r_script01.txt")`, donde el contenido de este fichero se encuentra en el apéndice C.2.

¹⁰`_dendrogram_se_part01.eps` y `_pco_se_part01.eps`.

¹¹Excepto si cierto número de observaciones consecutivas se encuentran dentro de esta zona y alejadas de la línea central.



(a) Dendrograma



(b) MDS

Fig. B.11: Representaciones gráficas de similitudes

tomar acción e investigar y ajustar el proceso, ya que seguramente existen causas especiales de variación presentes en el proceso.

Estas zonas pueden representarse con colores. Se suele utilizar el color verde en la primera zona para indicar la estabilidad del proceso, color amarillo en la segunda zona para indicar alerta en el proceso y color rojo en la tercera zona para indicar una alta variabilidad del proceso. En la figura B.12 se muestra la distribución normal acompañada de un gráfico de control con sus tres zonas.

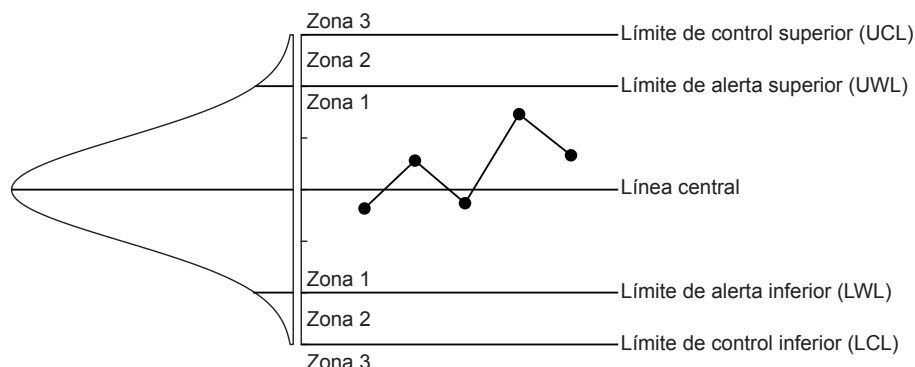


Fig. B.12: Gráfico de control con sus tres zonas

Como se observa en la figura B.12 la zona 1 se encuentra dentro de las primeras dos desviaciones ($\pm 2\sigma$). La segunda zona comienza a partir de la segunda desviación ($\pm 2\sigma$). Esta desviación representa los límites de alerta superior UWL (*Upper Warning Limit*, Límite de Alerta Superior) e inferior LWL (*Lower Warning Limit*, Límite de Alerta Inferior). Por último, la tercera zona comienza a partir de la tercera desviación ($\pm 3\sigma$).

En este tipo de gráficos no existe un procedimiento único que indique cuándo un proceso se considera fuera de control. No obstante, existen reglas que se utilizan como referencia para determinar si un proceso está fuera de control. Algunos de los conjuntos de reglas más conocidos son: las reglas de WECO (*Western Electric Company*) (WECO, 1956), las reglas de Wheeler (Wheeler, 1993) y las reglas de Nelson (Nelson, 1984). Por ejemplo, de acuerdo a las reglas de WECO (WECO, 1956) un proceso no está bajo control si se observa alguno de los siguientes patrones:

- Una observación se encuentra fuera de la tercera desviación (más allá de la zona 3).
- Dos de tres observaciones consecutivas (en un mismo lado de la línea central) se encuentran fuera de la segunda desviación (zona 2 o más allá).
- Cuatro de cinco observaciones sucesivas (en un mismo lado de la línea central) se encuentran fuera de la primera desviación (zona 1 o más allá).
- Ocho observaciones consecutivas que se encuentren en un mismo lado de la línea central dentro de la primera desviación (zona 1).

Existen diferentes tipos de gráficos de control (Oakland, 2003). Uno de los gráficos (gráfico de promedios) más usado en empresas del sector industrial consiste en obtener muestras en un periodo determinado y de una tamaño determinado de observaciones de alguna variable que se esté midiendo en el proceso. A cada muestra se calcula su promedio y se representa por una observación en el gráfico de control. Cuando no se dispone de numerosas muestras, se utiliza otro tipo de gráfico de control conocido como gráfico de

control individual o *i-Chart* (Oakland, 2003). En este gráfico cada observación representa un valor individual.

El tipo de gráfico de control usado para examinar la variabilidad en los tratamientos aplicados de las replicaciones de una partición es el gráfico de control individual, esto debido a que no se dispone de un gran número de observaciones como en el caso de las empresas del sector industrial y no se dispone de un número fijo de observaciones para constituir las muestras. En ISE cada experimento suele tener tamaños de muestra distintos.

Para realizar el gráfico de control individual, primero se calcula el promedio de las observaciones que se desean examinar con la función en (B.8),

$$\bar{x} = \frac{\sum_{i=1}^m x_i}{m} \quad (\text{B.8})$$

donde m es el número de observaciones y x_i es la observación i , en este caso el efecto i de uno de los tratamientos a examinar. Para calcular los límites de control superior UCL e inferior LCL se utiliza la función en (B.9),

$$\overline{MR} = \frac{\sum_{i=2}^m MR_i}{m - 1} \quad (\text{B.9})$$

donde \overline{MR} es el promedio de las diferencias absolutas de la observación x_i y su predecesora x_{i-1} i.e. $MR_i = |x_i - x_{i-1}|$. A partir de (B.9) se calcula la desviación para determinar los límites UCL, LCL, UWL y LWL. La desviación se calcula con la función en (B.10),

$$\sigma = \frac{\overline{MR}}{d_2} \quad (\text{B.10})$$

donde d_2 es la constante de Hartley (Hartley, 1942; Hartley y Pearson, 1942) conocida en SPC como constante de ajuste (d_2 o $(d)_n$) (Oakland, 2003; Montgomery, 1996).

Por tanto, los límites UCL, LCL, UWL y LWL se calculan a partir de la función en (B.11),

$$\begin{aligned} UCL &= \bar{x} + 3\sigma \\ LCL &= \bar{x} - 3\sigma \\ UWL &= \bar{x} + 2\sigma \\ LWL &= \bar{x} - 2\sigma \end{aligned} \quad (\text{B.11})$$

Para evaluar la tendencia de las observaciones en los gráficos de control individual, por ejemplo, Oakland (Oakland, 2003) recomienda seguir las siguientes reglas con el fin de determinar si el proceso está fuera de control:

- Cualquier observación fuera de los límites de la tercera desviación (zona 3).
- Dos de tres observaciones sucesivas fuera de la segunda desviación (zona 2).
- Cuatro de cinco observaciones sucesivas se encuentran fuera de la primera desviación (zona 1).
- Ocho observaciones consecutivas que se encuentren dentro de las primeras dos desviaciones pero alejadas de la línea central.

Un aspecto a considerar en este tipo de gráficos es que es relativamente menos sensible que el gráfico de control de promedios. En el gráfico de control de promedios al utilizar conjuntos de muestras donde cada muestra representa una observación, se tiene mayor precisión en las observaciones, esta precisión ayuda a detectar variaciones pequeñas en el proceso.

Otro aspecto a considerar en este tipo de gráficos es que existen diversas reglas que se pueden utilizar para evaluar las tendencias de las observaciones. Ya que en empresas del sector industrial cada proceso puede ser diferente, estas reglas algunas veces se modifican para no detener la producción por una observación que se encuentre fuera de control. En este sentido, se recomienda que el investigador quien ejecute el procedimiento para comparar replicas utilice como referencia la primera regla y el resto las ajuste a su contexto.

A continuación se ejemplifica las funciones antes descritas para generar los gráficos de control de dos tratamientos representados por las técnica A y B. El promedio de las observaciones de la técnica A en las replicas de una partición es,

$$\bar{x}_{tecA} = \frac{34.1 + 29.7 + 32.9}{3} = 32.2333$$

Aplicando la funciones en (B.9), (B.10) y (B.11) se obtiene,

$$\overline{MR}_{tecA} = \frac{|32.9 - 29.7| + |29.7 - 34.1|}{3 - 1} = \frac{3.2 + 4.4}{2} = 3.8$$

$$\sigma = \frac{3.8}{1.128} = 3.3688$$

$$UCL_{tecA} = 32.2333 + 3(3.3688) = 42.3397$$

$$LCL_{tecA} = 32.2333 - 3(3.3688) = 22.1270$$

$$UWL_{tecA} = 32.2333 + 2(3.3688) = 38.9709$$

$$LWL_{tecA} = 32.2333 - 2(3.3688) = 25.4957$$

para el caso de las observaciones de la técnica B en las replicas de la partición se obtiene,

$$\bar{x}_{tecA} = \frac{28.1 + 22.6 + 37.5}{3} = 29.4$$

$$\overline{MR}_{tecA} = \frac{|37.5 - 22.6| + |22.6 - 28.1|}{3 - 1} = \frac{14.9 + 5.5}{2} = 10.2$$

$$\sigma = \frac{10.2}{1.128} = 9.0426$$

$$UCL_{tecA} = 29.4 + 3(9.0426) = 56.5277$$

$$LCL_{tecA} = 29.4 - 3(9.0426) = 2.2723$$

$$UWL_{tecA} = 29.4 + 2(9.0426) = 47.4851$$

$$LWL_{tecA} = 29.4 - 2(9.0426) = 11.3149$$

con los resultados de las funciones anteriores se construye el gráfico de control. En la figura B.13 se muestran los gráficos de control resultantes.

B.8 Meta-análisis

Meta-análisis es un término que se ha acuñado a una familia de procedimientos estadísticos (Pearson, 1904; Tippet, 1931; Fisher, 1932; Cochran, 1937, 1954; Stouffer et al., 1949; Jones y Fiske, 1953; Winer, 1971; Glass, 1976; Rosenthal, 1978; George, 1977; Hedges y Olkin, 1985; Higgins y Thompson, 2002; Hunter y Schmidt, 2004) usados para combinar de forma cuantitativa los resultados de estudios, replicaciones o experimentos similares.

De manera general un meta-análisis consiste en: 1) estimar el tamaño del efecto de un par de tratamientos por cada uno de los experimentos que se desean examinar. 2) estimar un tamaño de efecto global a partir de los tamaños de efecto estimados individualmente, y 3) estimar el nivel de heterogeneidad del conjunto de tamaños de efecto observados. Los resultados de estos cálculos suelen representarse en un diagrama de bosque (Forest plot en inglés).

B.8.1 Estimación de Tamaños de Efecto Individuales

El tamaño del efecto es un índice que se utiliza para medir las diferencias entre dos variables, en este caso un par de tratamientos. El tamaño del efecto permite conocer qué tanto es mejor un tratamiento frente a otro. Para calcular el tamaño del efecto de cada replicación de la partición seleccionada se utiliza la técnica de Glass (Glass, 1976) conocida como diferencia de medias ponderada WMD (*Weighted Mean Difference*). Esta técnica ha sido ampliamente usada en disciplinas afines a la medicina y a las ciencias sociales (Hunter y Schmidt, 2004). En IS esta técnica también ha sido usada por varios investigadores (Hayes, 1999; Miller, 2000; Hannay et al., 2009). Como primer paso de esta técnica se calcula el tamaño del efecto de dos tratamientos con la función en (B.12) (Borenstein et al., 2009),

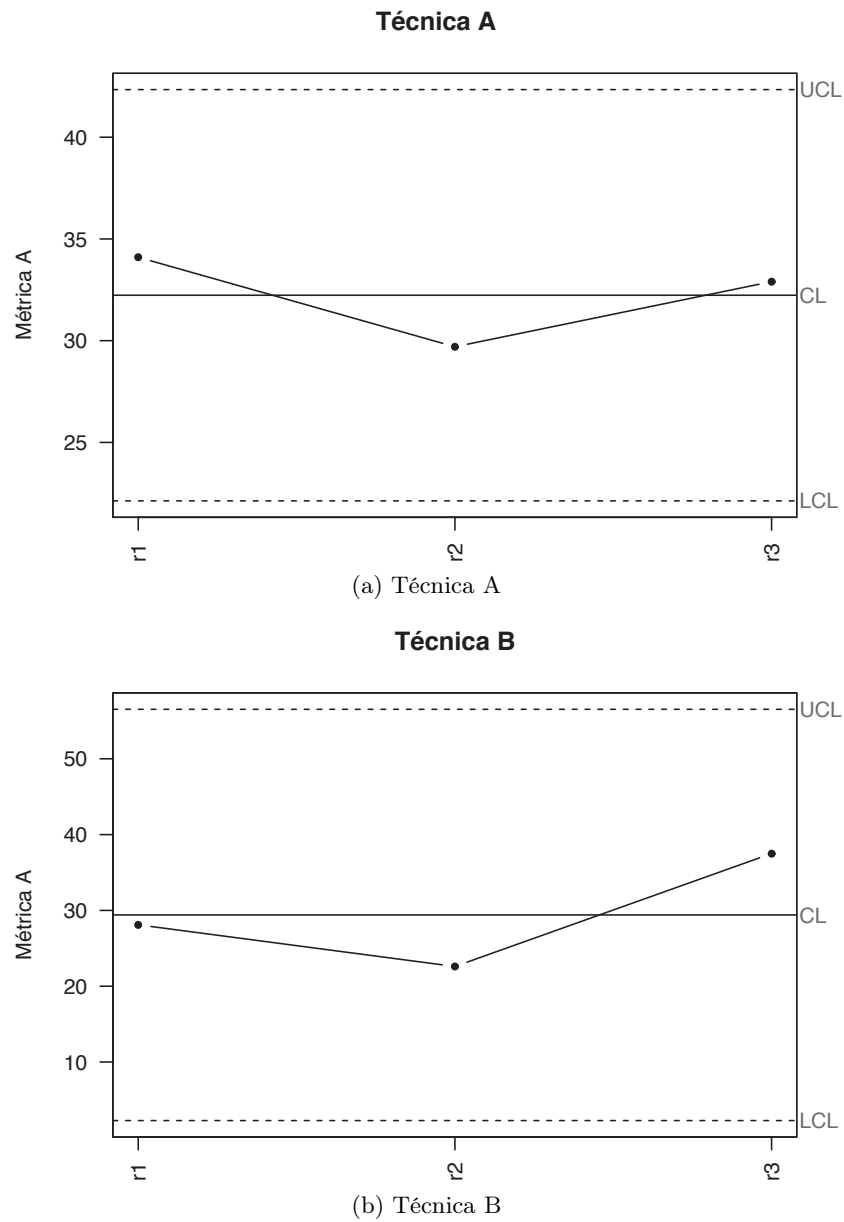


Fig. B.13: Gráficos de control resultantes

$$d = \frac{\bar{X}_1 - \bar{X}_2}{S_{within}} \quad (\text{B.12})$$

donde d es el tamaño del efecto, \bar{X}_1 y \bar{X}_2 son los promedios muestrales de los tratamientos¹² y S_{within} es la desviación estándar (intra-grupos) agrupada de ambos promedios muestrales y se calcula a partir de la función en (B.13) (Borenstein et al., 2009),

$$S_{within} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \quad (\text{B.13})$$

donde S_{within} es la desviación estándar agrupada de ambos promedios muestrales, S_1 y S_2 son las desviaciones estándar de los promedios muestrales de los tratamientos \bar{X}_1 y \bar{X}_2 , n_1 y n_2 son los tamaños de las muestras usadas en los promedios muestrales de los tratamientos \bar{X}_1 y \bar{X}_2 .

La varianza de d se estima a partir de la función descrita en (B.14) (Borenstein et al., 2009),

$$V_d = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)} \quad (\text{B.14})$$

el resultado del índice o coeficiente d (ver función B.12) puede ser:

1. Positivo, que indica que \bar{X}_1 es mejor que \bar{X}_2 , o que el grupo experimental se comporta mejor que el grupo de control.
2. Negativo, que indica que \bar{X}_2 es mejor que \bar{X}_1 , o que el grupo de control se comporta mejor que el grupo experimental.
3. Cero, que indica que no hay diferencias entre \bar{X}_1 y \bar{X}_2 , o que no hay diferencias entre el grupo experimental y el grupo de control.

A su vez el tamaño de efecto observado d puede ser clasificado como efecto de tamaño pequeño, mediano o grande para los valores absolutos de d comprendidos entre 0.2 y 0.3, 0.5, y 0.8 respectivamente (Cohen, 1988).

La función descrita en (B.12) para calcular el coeficiente d (Glass, 1976) fue mejorada por Hedges y Olkin (Hedges y Olkin, 1985). Hedges y Olkin añadieron un factor de corrección J para mejorar la precisión de los resultados cuando en los experimentos se cuenta con un tamaño de muestra pequeño (especialmente por debajo de 10), como en el caso de los experimentos que se llevan a cabo en IS donde el tamaño de la muestra suele ser pequeño. El factor de corrección J se calcula a partir de (B.15) (Borenstein et al., 2009),

¹²Estos estimadores estadísticos también suelen identificarse como el promedio de la muestra que contiene la intervención o tratamiento a examinar (grupo experimental) \bar{X}_E , y por el promedio de la muestra que se utiliza como control (grupo de control) \bar{X}_C .

$$J = 1 - \frac{3}{4df - 1} \quad (\text{B.15})$$

donde J es el factor de corrección, y df son los grados de libertad usados para estimar S_{within} , los cuales son $n_1 + n_2 - 2$.

A partir del coeficiente d en (B.12) y el factor de corrección J en (B.15), se obtiene el coeficiente g en (B.16) (Borenstein et al., 2009) que contiene el tamaño de efecto ajustado,

$$g = J \times d \quad (\text{B.16})$$

la varianza de g se calcula a partir de la función en (B.17) (Borenstein et al., 2009),

$$V_g = J^2 \times V_d \quad (\text{B.17})$$

y el error estándar de g se obtiene con la función en (B.18) (Borenstein et al., 2009),

$$SE_g = \sqrt{V_g} \quad (\text{B.18})$$

ya que el coeficiente resultante en (B.16) es un estimador del tamaño del efecto verdadero i.e. del parámetro poblacional, es necesario calcular el intervalo de confianza en el que puede estar el parámetro poblacional. Para esto se utiliza la función en (B.19),

$$g - Z_{\alpha/2}SE_g \leq g \leq g + Z_{\alpha/2}SE_g \quad (\text{B.19})$$

donde g es el tamaño del efecto ajustado, Z es el número de desviaciones estándar que separan, a un nivel de significación determinado, el tamaño de efecto estimado de los límites, 1.96 ($\alpha = 0.05$), y SE_g es el error estándar de g .

A partir de la tabla de resultados experimentales generada en el sub-proceso *análisis de resultados experimentales* se calcula por cada replicación el tamaño de efecto estimado así como su intervalo de confianza. En la tabla B.20 se muestra un ejemplo con los resultados experimentales de tres replicaciones que conforman una partición. En este ejemplo se tienen los resultados experimentales de tres replicaciones que evalúan la efectividad representada con la métrica A de las técnicas A y B.

Tomando como referencia los resultados experimentales de la replicación *r1*, a continuación se calcula el tamaño de efecto y sus intervalos de confianza de acuerdo a las funciones antes descritas. En la tabla B.21 se muestran los resultados de la replicación *r1*.

Técnica A				Técnica B		
Id	Tamaño	Métrica A	Desv. est.	Tamaño	Métrica A	Desv. est.
exp.	(n_1)	($\bar{\chi}_1$)	(s_1)	(n_2)	($\bar{\chi}_2$)	(s_2)
r1	45	34.1	12.7	45	28.1	11.5
r2	39	29.7	22.1	39	22.6	25.8
r3	43	32.9	17.5	43	37.5	21.2

Tabla B.20: Resultados experimentales de tres replicaciones que pertenecen a una partición

Técnica A				Técnica B		
Id	Tamaño	Métrica A	Desv. est.	Tamaño	Métrica A	Desv. est.
exp.	(n_1)	($\bar{\chi}_1$)	(s_1)	(n_2)	($\bar{\chi}_2$)	(s_2)
r1	45	34.1	12.7	45	28.1	11.5

Tabla B.21: Resultados de la replicación $r1$

Para calcular el tamaño del efecto estimado g que contiene el factor de corrección, primero es necesario calcular la desviación estándar agrupada, el tamaño del efecto estimado d , así como su varianza V_d .

La desviación estándar agrupada se obtiene a partir de (B.13) (Borenstein et al., 2009),

$$S_{within} = \sqrt{\frac{(45-1)12.7^2 + (45-1)11.5^2}{45+45-2}} = 12.1149$$

Una vez obtenida S_{within} se calcula el tamaño del efecto estimado d de acuerdo a la función en (B.12) (Borenstein et al., 2009),

$$d = \frac{34.1 - 28.1}{12.1149} = 0.4953$$

Con las función en (B.14) se obtiene la varianza de d (Borenstein et al., 2009),

$$V_d = \frac{45+45}{45 \times 45} + \frac{0.4953^2}{2(45+45)} = 0.0458$$

Para obtener el tamaño del efecto estimado con el ajuste se aplica la función en (B.15) para obtener J (Borenstein et al., 2009),

$$J = 1 - \frac{3}{4(45+45-2)-1} = 0.9915$$

por lo que

$$g = 0.9915 \times 0.4953 = 0.4910$$

La varianza y el error estándar de g se calculan a partir de (B.17) y (B.18) (Borenstein et al., 2009),

$$V_g = 0.9915^2 \times 0.0458 = 0.0450$$

y el error estándar es (Borenstein et al., 2009),

$$SE_g = \sqrt{0.0450} = 0.2122$$

De acuerdo al resultado de g se observa un tamaño de efecto positivo indicando que en la replicación $r1$, la técnica A es más efectiva que la técnica B. Respecto a la clasificación de tamaños de efecto de Cohen (Cohen, 1988), el tamaño de efecto resultante se considera mediano. Sin embargo al obtener un tamaño de efecto estimado se requiere calcular un intervalo de confianza en el que se estime pueda estar el tamaño de efecto verdadero. En este caso se realizan los cálculos para estimar el parámetro poblacional en un intervalo de confianza del 95% ($\alpha = 0.05$). De acuerdo a (B.19) se obtienen los siguientes intervalos,

$$Lim_{inf} = 0.4910 - 1.96 \times 0.2122 = 0.0751$$

$$Lim_{sup} = 0.4910 + 1.96 \times 0.2122 = 0.9069$$

como se observa en los límites estimados en (B.21), el tamaño de efecto verdadero puede estar entre los límites [0.0751; 0.9069]. Ya que los intervalos no solapan con el valor 0.0 el tamaño de efecto estimado indica que la técnica A es significativamente más efectiva que la técnica B. Usando las funciones antes descritas, en la tabla B.22 se muestran los tamaños de efecto estimados así como sus intervalos de confianza a un nivel del 95%.

Técnica A				Técnica B					
Id exp.	Tamaño (n1)	Métrica A ($\bar{y}1$)	Desv. est. (s1)	Tamaño (n2)	Métrica A ($\bar{y}2$)	Desv. est. (s2)	Tam. efecto est. (g)	Lim. inf. 95% I.C.	Lim. sup. 95% I.C.
r1	45	34.1	12.7	45	28.1	11.5	0.4910	0.0751	0.9069
r2	39	29.7	22.1	39	22.6	25.8	0.2926	-0.1492	0.7345
r3	43	32.9	17.5	43	37.5	21.2	-0.2345	-0.6549	0.1859

Tabla B.22: Tamaños de efecto estimados e intervalos de confianza

B.8.2 Estimación de Tamaño de Efecto Global

Una vez estimados los tamaños de los efectos y sus intervalos de confianza, a continuación se combinan estos efectos para obtener un efecto global de las replicaciones que componen alguna partición. El efecto global es un promedio ponderado del conjunto de tamaños de efecto estimados. A cada tamaño de efecto estimado se le asigna un determinado nivel de ponderación o peso de acuerdo al nivel de precisión que éste tiene. Los tamaños de efecto que tienen mayor nivel de precisión se les asigna mayor peso.

La precisión en el tamaño de efecto está en función, principalmente, del tamaño de muestra utilizado en el experimento. A mayor tamaño de muestra los intervalos de confianza del tamaño de efecto estimado se reducen, por lo que el tamaño de efecto estimado

se aproxima mucho más al tamaño de efecto verdadero. Por otra parte, a menor tamaño de muestra los intervalos de confianza del tamaño de efecto estimado se amplían, por lo que es más probable que el tamaño de efecto estimado se aproxime menos al tamaño de efecto verdadero.

Dos de los modelos estadísticos más utilizados para calcular el efecto global son, el modelo de efecto fijo y el modelo de efectos aleatorios (Hedges y Olkin, 1985; Borenstein et al., 2009). La elección de un modelo u otro depende del supuesto que se tenga acerca de la distribución de los tamaños de los efectos que provienen del conjunto de replicaciones, en este caso de la partición. En el modelo de efecto fijo, se asume que todas las replicaciones en este caso de la partición comparten un mismo tamaño de efecto verdadero i.e. se asume un tamaño de efecto verdadero que es fijo y subyace en todas las replicaciones, por lo que las diferencias observadas en los tamaños de efecto estimados son producto del error muestral y no producto de la influencia de otros factores o condiciones experimentales. En este modelo, el efecto global es un estimador del tamaño de efecto verdadero que es común en las replicaciones de la partición. Por el contrario, en el modelo de efectos aleatorios se asume que el tamaño de efecto verdadero puede variar de un experimento a otro i.e. en las replicaciones de la partición pueden existir factores o condiciones que influyen en el tamaño del efecto verdadero. En este modelo se asume que los tamaños de efecto verdaderos se distribuyen de acuerdo a la distribución normal, y el efecto global es entonces un estimador del promedio de la distribución de tamaños de efecto verdaderos.

Como se discute en el capítulo 5, no es posible tener replicaciones exactas de algún experimento porque siempre existen condiciones o factores que varían entre la replicación y el experimento base. Al haber en las replicaciones cambios en los elementos de la situación experimental tanto deliberados como no deliberados los tamaños de efecto verdaderos varían entre las replicaciones. Ante a esta situación, el modelo estadístico que se utiliza para calcular el efecto global en las replicaciones de una partición es el modelo de efectos aleatorios. Para calcular el efecto global con el modelo de efectos aleatorios se utiliza la función descrita en (B.20) (Borenstein et al., 2009),

$$M^* = \frac{\sum_{i=1}^k W_i^* Y_i}{\sum_{i=1}^k W_i^*} \quad (\text{B.20})$$

donde M^* es el efecto global estimado, W_i es la inversa de la varianza intra-estudio del experimento $_i$ más la varianza inter-estudios, i.e. representa el peso asignado al experimento $_i$, Y_i es el efecto observado en el experimento $_i$ i.e. el tamaño de efecto estimado g .

Para obtener el efecto global estimado en (B.20) es necesario calcular antes las funciones descritas en (B.21), (B.22), (B.23), (B.24), (B.25) (Borenstein et al., 2009),

$$W_i^* = \frac{1}{V_{y_i}^*} \quad (\text{B.21})$$

donde W_i^* es el peso asignado a cada experimento, $V_{y_i}^*$ es la varianza intra-estudio de cada experimento más la varianza inter-estudios y se obtiene a partir de la función en (B.22) (Borenstein et al., 2009),

$$V_{y_i}^* = V_{y_i} + T^2 \quad (\text{B.22})$$

donde V_{y_i} es la varianza intra-estudio del experimento_{*i*} y T^2 es la varianza inter-estudios del conjunto de replicaciones que componen la partición, y se calcula con la función en (B.23) (Borenstein et al., 2009),

$$T^2 = \frac{Q - df}{C} \quad (\text{B.23})$$

donde T^2 es la varianza inter-estudios, Q es la suma de las desviaciones cuadradas ponderadas WSS (*Weighted Sum of Squares*), df son los grados de libertad $k - 1$ donde k es el número de experimentos, y C es una función complementaria. Ya que la varianza en los tamaños de efecto verdaderos τ^2 nunca puede ser menor que cero, el estimado de la varianza T^2 tampoco puede ser menor que cero, en el caso de obtener un valor menor a cero se debe al error muestral. En esta situación se asigna el valor 0 al estimador T^2 (Borenstein et al., 2009).

Para obtener Q y C se utilizan las funciones en (B.24), (B.25) (Borenstein et al., 2009),

$$Q = \sum_{i=1}^k W_i Y_i^2 - \frac{\left(\sum_{i=1}^k W_i Y_i\right)^2}{\sum_{i=1}^k W_i} \quad (\text{B.24})$$

$$C = \sum W_i - \frac{\sum W_i^2}{\sum W_i} \quad (\text{B.25})$$

donde $W_i = \frac{1}{Y_i}$ que es la inversa de la varianza del efecto observado Y del experimento_{*i*} i.e. el tamaño de efecto estimado g .

La varianza del efecto global estimado M^* así como su error estándar estimado se calculan de acuerdo a las funciones en (B.26) y (B.27) (Borenstein et al., 2009),

$$V_{M^*} = \frac{1}{\sum_{i=1}^k W_i^*} \quad (\text{B.26})$$

$$SE_{M^*} = \sqrt{V_{M^*}} \quad (\text{B.27})$$

una vez estimado el tamaño de efecto global, se utiliza la función en (B.29) para estimar el intervalo de confianza (Borenstein et al., 2009),

$$M^* - Z_{\alpha/2}SE_{M^*} \leq M^* \leq M^* + Z_{\alpha/2}SE_{M^*} \quad (\text{B.28})$$

a partir de los tamaños de efecto estimados que se muestran en la tabla B.22, a continuación se calcula el tamaño de efecto global utilizando las funciones antes descritas. Primero se calcula W , WY , WY^2 y W^2 para estimar Q y C . En la tabla B.23 se muestran los cálculos de W , WY , WY^2 y W^2 .

Id exp.	Efecto Y (g)	Intra-Varianza $V_y(V_g)$	Peso (W)	WY	WY ²	W ²
r1	0.4910	0.0450	22.2087	10.9050	5.3547	493.2255
r2	0.2926	0.0508	19.6771	5.7584	1.6852	387.1872
r3	-0.2345	0.0460	21.7381	-5.0982	1.1957	472.5466
		Suma	63.6239	11.5652	8.2355	1352.9593

Tabla B.23: Resultados de W , WY , WY^2 y W^2 para calcular Q y C

A partir de los resultados de la tabla B.23 se aplican las funciones (B.24) y (B.25) (Borenstein et al., 2009),

$$Q = 8.2355 - \frac{(11.5652)^2}{63.6239} = 6.1332$$

$$C = 63.6239 - \frac{1352.9593}{63.6239} = 42.3589$$

con los resultados de Q y C se calcula la varianza inter-estudios aplicando la función en (B.23) (Borenstein et al., 2009),

$$T^2 = \frac{6.1332 - (3 - 1)}{42.3589} = 0.0976$$

A continuación se calcula la varianza total y el peso. En la tabla B.24 se presentan los resultados de estos cálculos que se utilizan para estimar el tamaño del efecto global.

Id exp.	Efecto Y (g)	Intra-Varianza $V_y(V_g)$	Inter-Varianza (T^2)	Varianza total (V^*_y)	Peso (W^*)	W*Y
r1	0.4910	0.0450	0.0976	0.1426	7.0124	3.4433
r2	0.2926	0.0508	0.0976	0.1484	6.7387	1.9720
r3	-0.2345	0.0460	0.0976	0.1436	6.9648	-1.6334
				Suma	20.7160	3.7819

Tabla B.24: Resultados de varianza total y peso en replicaciones de una partición

El tamaño de efecto global se estima a partir de la función en (B.20) (Borenstein et al., 2009),

$$M^* = \frac{3.7819}{20.7160} = 0.1826$$

la varianza del tamaño de efecto global así como su error estándar se calculan a partir de (B.26) y (B.27) (Borenstein et al., 2009),

$$V_{M^*} = \frac{1}{20.7160} = 0.0483$$

$$SE_{M^*} = \sqrt{0.0483} = 0.2197$$

para calcular el intervalo de confianza del tamaño de efecto global estimado se aplica la función en (B.29), donde se utiliza un nivel de confianza del 95% ($\alpha = 0.05$),

$$Lim_{inf} = 0.1826 - 1.96 \times 0.2197 = -0.2481$$

$$Lim_{sup} = 0.1826 + 1.96 \times 0.2197 = 0.6132$$

de acuerdo al tamaño de efecto global estimado, se observa un efecto positivo de tamaño pequeño a favor de la técnica A. Al observar los límites estimados, el tamaño de efecto verdadero global puede estar entre los límites $[-0.2481; 0.6132]$ por lo que se puede afirmar que la técnica A es ligeramente más efectiva que la técnica B.

En la tabla B.25 se muestran los resultados experimentales, los tamaños de los efectos estimados, los intervalos de confianza y el peso de las replicaciones de la partición, también se muestra el tamaño de efecto global estimado así como sus intervalos de confianza.

Los resultados de un meta-análisis suelen representarse de manera gráfica a través de un diagrama de bosque. En la figura B.14 se muestra el diagrama de bosque con los tamaños de efecto estimados así como el tamaño de efecto global estimado.

Id exp.	Téc. A			Téc. B			Tam. ef. est. (g)	Lim. inf. 95% I.C.	Lim. sup. 95% I.C.	Peso	
	Tam. (n _i)	Métrica A ($\bar{\chi}_1$)	Desv. est. (s _i)	Tam. (n ₂)	Métrica A ($\bar{\chi}_2$)	Desv. est. (s ₂)				Peso (W*)	Relativo
r1	45	34.1	12.7	45	28.1	11.5	0.4910	0.0751	0.9069	7.0124	33.9%
r2	39	29.7	22.1	39	22.6	25.8	0.2926	-0.1492	0.7345	6.7387	32.5%
r3	43	32.9	17.5	43	37.5	21.2	-0.2345	-0.6549	0.1859	6.9648	33.6%
Tam. efecto global est. (M*)							0.1826	-0.2481	0.6132	20.7160	100.0%

Tabla B.25: Resultados experimentales, tamaños de efecto, límites, pesos y efecto global de las replicaciones de una partición

En el diagrama de la figura B.14 las primeras tres líneas horizontales representan las replicaciones de la partición: $r1$, $r2$ y $r3$. El tamaño del efecto estimado en cada replicación se representa por un cuadrado, los intervalos de confianza de los tamaños de efecto estimados representan las líneas horizontales, el peso asignado a cada replicación se representa por el tamaño del cuadrado, entre mayor sea el tamaño mayor es el peso que se asigna a la replicación. El efecto global estimado se representa por el rombo que se encuentra en la parte inferior del diagrama, su anchura representa los intervalos de confianza estimados.

Para concluir el meta-análisis de las replicaciones de la partición, resta evaluar el nivel de variabilidad o heterogeneidad del conjunto de tamaños de efecto estimados. La variabilidad del conjunto de efectos observados en las replicaciones de la partición se evalúa a través de una prueba de heterogeneidad (Cochran, 1954) que utiliza el estadístico Q así como por los resultados los estadísticos T^2 e I^2 (Higgins y Thompson, 2002).

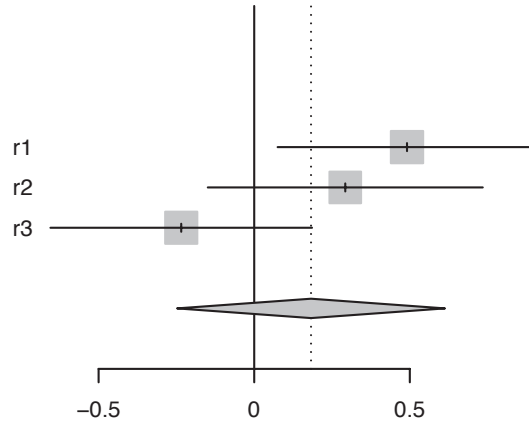


Fig. B.14: Diagrama de bosque de los efectos estimados en tres replicaciones de una partición

B.8.3 Heterogeneidad

La prueba de heterogeneidad utiliza el índice Q para evaluar si existe heterogeneidad estadísticamente significativa en un conjunto de tamaños de efecto estimados. Esta prueba parte de la hipótesis nula de la existencia de un tamaño de efecto que es común en los tamaños de efecto de un conjunto de experimentos. Bajo la hipótesis nula Q sigue una distribución χ^2 (chi-cuadrada) con $k - 1$ grados de libertad donde k es el número de experimentos. El no rechazar la hipótesis nula se asume que los efectos estimados son homogéneos i.e. no hay variabilidad en los resultados. Por otra parte, el rechazar la hipótesis nula se asume una heterogeneidad en los efectos estimados. El valor α que se suele utilizar para esta prueba es de 0.05, por lo que un valor p menor que α conlleva a rechazar la hipótesis nula.

No obstante este test tiene una potencia baja para detectar una verdadera heterogeneidad cuando se tiene un conjunto reducido de experimentos (Huedo-Medina et al., 2006; Borenstein et al., 2009). Debido a esta situación se utiliza también el estadístico T^2 descrito anteriormente, que consiste en estimar la varianza inter-estudios de un conjunto de experimentos. El otro estadístico que también se utiliza es el índice I^2 (Higgins y Thompson, 2002) que se interpreta como el porcentaje de variabilidad total en un conjunto de tamaños de efecto estimados, que se debe a una verdadera heterogeneidad. Los autores proponen tres valores de referencia para este índice, sugieren que valores alrededor de 25%, 50% y 75% representan respectivamente un nivel de heterogeneidad bajo, medio y alto. El índice I^2 (Higgins y Thompson, 2002) se calcula a partir de la función en (B.29) (Higgins y Thompson, 2002; Borenstein et al., 2009),

$$I^2 = \left(\frac{Q - df}{Q} \right) \times 100\% \quad (\text{B.29})$$

donde I^2 es el porcentaje estimado de heterogeneidad verdadera, Q es la suma de las desviaciones cuadradas ponderadas observada WSS, df son los grados de libertad $k - 1$ donde k es el número de experimentos y representa la suma de las desviaciones cuadradas ponderadas esperada bajo el supuesto que los estudios comparten un efecto común, la

diferencia $Q - df$ refleja el exceso de variación. En caso que $Q < df$ sea menor que cero se asigna el valor 0 al numerador. Esta situación se presenta cuando la variación observada Q es menor a la variación esperada df .

Usando como ejemplo la información de la tabla B.25, se aplica la prueba de heterogeneidad con un nivel α de 0.05, así como la funciones en (B.23), (B.24) y (B.29) (Borenstein et al., 2009) donde se obtiene,

$$Q = 6.1332$$

$$df = 2$$

$$p = 0.0466$$

$$T^2 = 0.0976$$

$$I^2 = \left(\frac{6.1332 - 2}{6.1332} \right) \times 100\% = \left(\frac{4.1332}{6.1332} \right) \times 100\% = 67.3907\%$$

de acuerdo a los resultados obtenidos, se observa que los tamaños de efecto estimados en las replicaciones de la partición son heterogéneos¹³, se observa que la replicación $r3$ contiene un efecto negativo que a diferencia del resto de replicaciones indica que la técnica B es ligeramente más efectiva¹⁴ que la técnica A.

B.9 Automatización de Cálculos para el Análisis de Resultados

Con la finalidad de reducir el esfuerzo que conlleva realizar los gráficos de control y el meta-análisis, se utiliza un fichero de texto que contiene varias funciones escritas en el lenguaje R (R Dev. Core Team, 2011). Parte de estas funciones utilizan el paquete *qcc* (Scrucca, 2004) para realizar los gráficos de control así como el paquete *meta* (Schwarzer, 2010) para realizar los cálculos referentes al meta-análisis. De la misma manera que se carga en el entorno R (R Dev. Core Team, 2011) el fichero de texto para realizar el análisis de conglomerados jerárquico y el análisis de escalado multidimensional, se carga el fichero de texto encargado de realizar el meta-análisis y los gráficos de control de las replicaciones de la partición seleccionada. Para cargar el fichero se utiliza la función `source()`¹⁵. En este fichero cargan los resultados experimentales generados en la tabla de resultados experimentales. La tabla con los resultados experimentales se guarda como un fichero de texto delimitado por tabuladores¹⁶. Tras cargar el fichero `script02.txt` y realizar varias funciones, se generan los ficheros `_meta_part01.eps` que contiene el diagrama de bosque con los resultados del meta-análisis así como los ficheros `_cc_a_part01.eps` y `_cc_b_part01.eps` que contienen los gráficos de control de los dos tratamientos de las replicaciones de la partición seleccionada. En la figura B.15 se muestra un ejemplo con

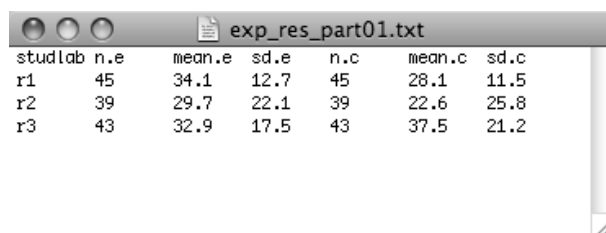
¹³El resultado de la prueba de heterogeneidad es significativo, de acuerdo al valor establecido $\alpha = 0.05$, $p = 0.0466$; por otra parte los valores de T^2 e I^2 son elevados.

¹⁴Aunque la diferencia no es significativa ya que el tamaño del efecto real puede ser 0.

¹⁵`source("part01/res_comp/r_script02.txt")` donde el contenido de este fichero se encuentra disponible en el apéndice C.3.

¹⁶La estructura de este fichero es muy similar a la tabla del formulario PCR-7, en lugar de usar los encabezados de este formulario se sustituyen por los encabezados: *studlab*, *n.e.*, *mean.e.*, *sd.e.*, *n.c.*, *mean.c* y *sd.c*; cabe recordar que las columnas deben estar delimitadas por tabuladores.

la estructura del fichero de texto `exp_res_part01.txt` con resultados experimentales de una partición.



studlab	n.e	mean.e	sd.e	n.c	mean.c	sd.c
r1	45	34.1	12.7	45	28.1	11.5
r2	39	29.7	22.1	39	22.6	25.8
r3	43	32.9	17.5	43	37.5	21.2

Fig. B.15: Fichero de texto con la tabla de resultados experimentales de una partición

Por otra parte, en las figuras B.16 y B.17 se muestran los diagramas generados tras cargar el fichero `r_script02.txt` en el entorno R (R Dev. Core Team, 2011).

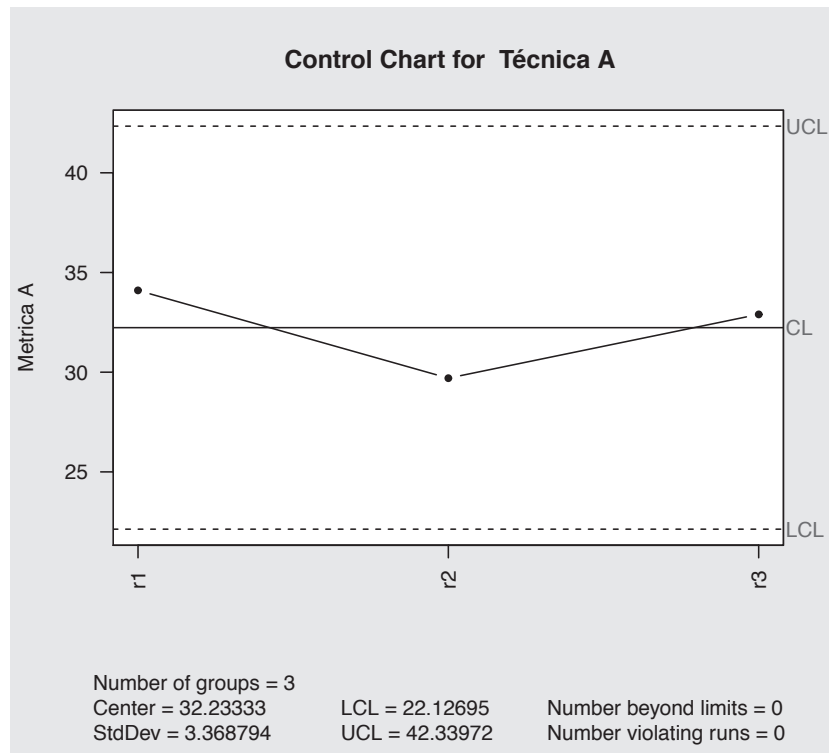
B.10 Análisis de Causas

Como herramienta para realizar el análisis de causas de resultados inconsistentes se usan diagramas de Ishikawa (Ishikawa, 1990) conocidos también como diagramas de causa y efecto. El diagrama de Ishikawa (Ishikawa, 1990) es una representación gráfica que ayuda a ordenar de forma concentrada las causas que se cree influyen en un determinado efecto. Este tipo de diagrama es comúnmente usado en empresas del sector industrial y de servicios, donde se utiliza para facilitar el análisis de problemas y sus soluciones relacionados con la calidad de los procesos, productos o servicios. En esta actividad se utiliza este tipo de diagrama para identificar posibles elementos de la situación experimental que se cree influyen en los resultados experimentales. En la figura B.18 se muestra una adaptación de este diagrama para llevar a cabo el análisis de causas en una partición donde se obtienen resultados inconsistentes.

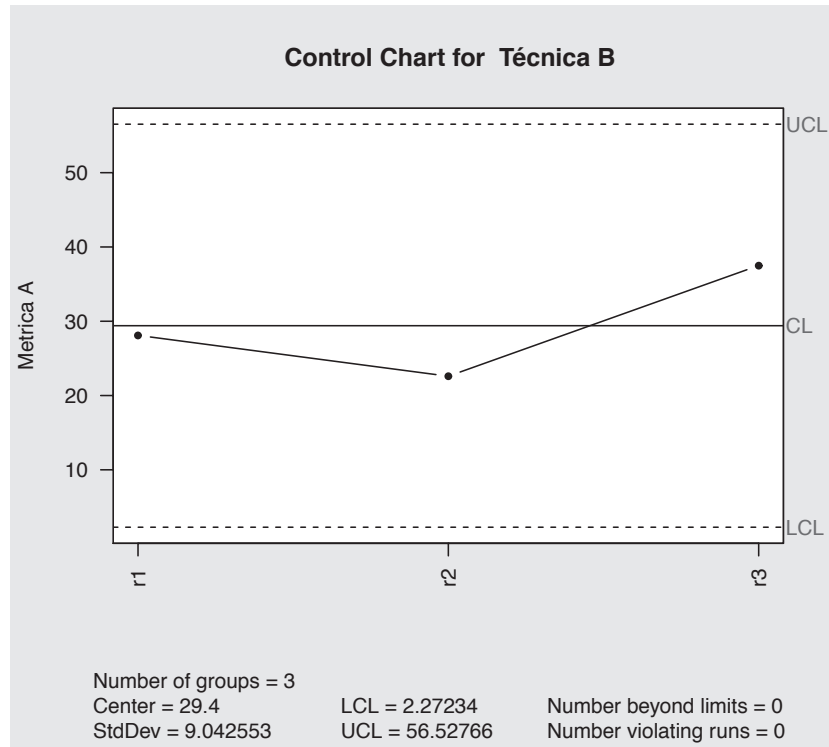
La parte izquierda del diagrama contiene las posibles causas, mientras que la parte derecha el efecto observado. El efecto puede representarse con los resultados del meta-análisis y/o resultados de los gráficos de control. En esta adaptación del diagrama de Ishikawa (Ishikawa, 1990) las causas principales representan las secciones que conforman la situación experimental. Las causas principales pueden contener causas más detalladas. Por ejemplo, en el diagrama de la figura B.18 cada causa principal contiene los elementos de las diferentes secciones de la situación experimental. En la figura B.19 se muestra un nivel más detallado en donde se describen las causas a nivel de sub-elementos y elementos del protocolo de la situación experimental.

Una vez que se tiene uno o varios diagramas de Ishikawa (Ishikawa, 1990) de alguna partición, se realiza el análisis de causas discutiendo y marcando los cambios identificados en los elementos de la situación experimental que se cree influyen en los resultados. En la figura B.20 se muestra un diagrama con el número de cambios identificados en los roles de los experimentadores de las replicaciones de una partición.

Una vez identificados los cambios se discuten los elementos que se creen influyen en los resultados experimentales de la partición. Estos cambios se registran para estudiarlos en profundidad en replicaciones posteriores.



(a) Técnica A



(b) Técnica B

Fig. B.16: Gráficos de control resultantes

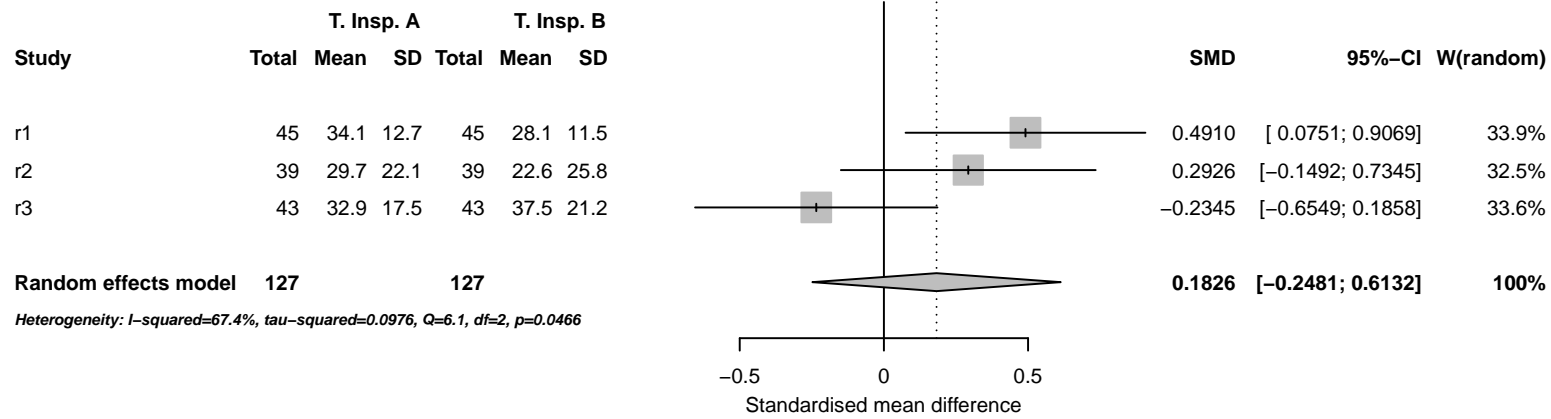


Fig. B.17: Meta-análisis de la partición seleccionada

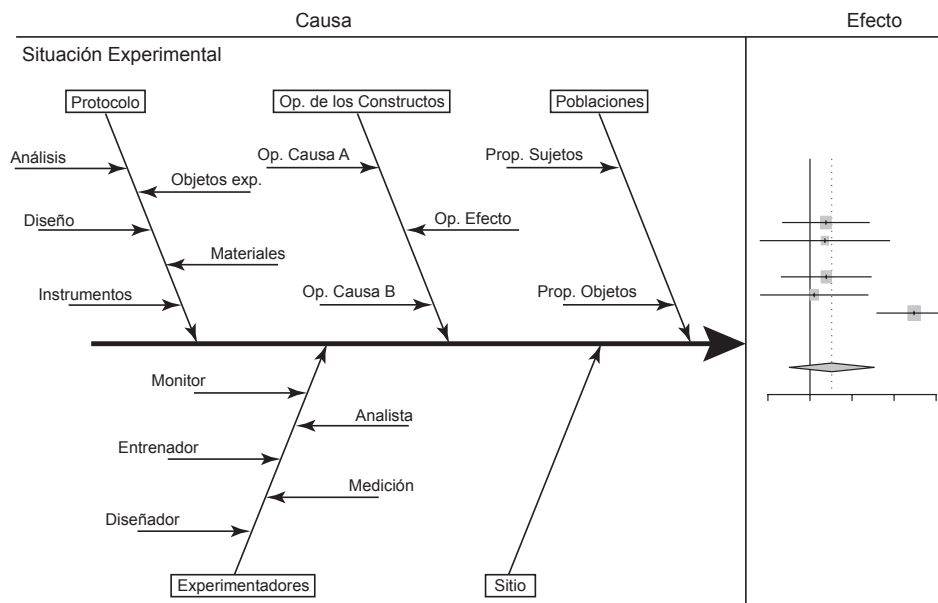


Fig. B.18: Diagrama de Ishikawa usado en el análisis de causas

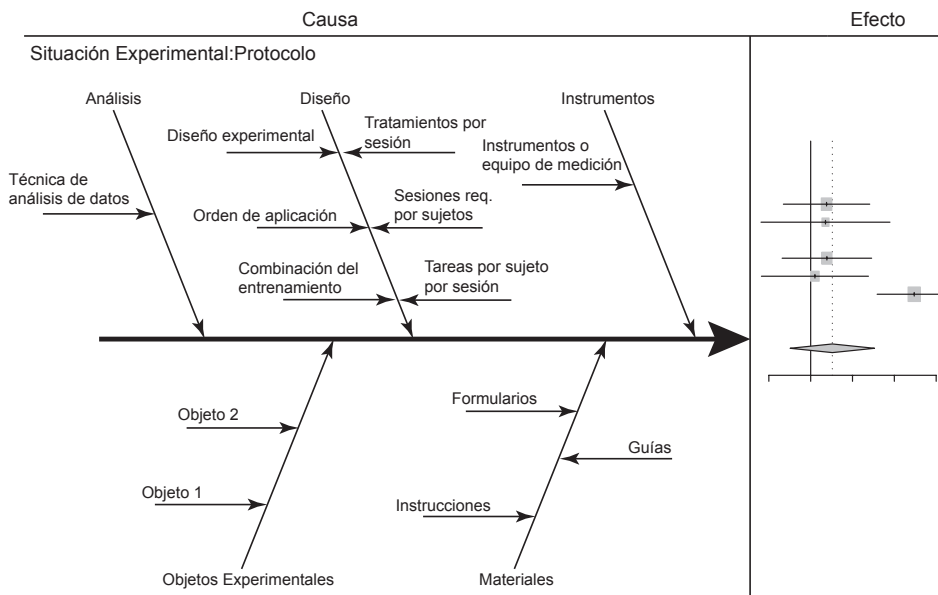


Fig. B.19: Diagrama de Ishikawa para el protocolo experimental

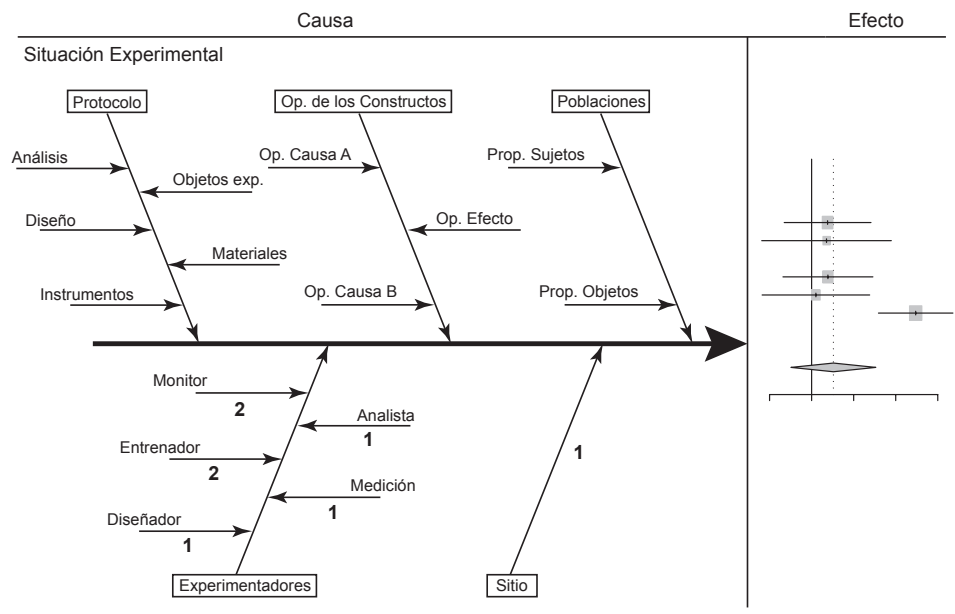


Fig. B.20: Diagrama de Ishikawa con número de cambios identificados

Apéndice C

Códigos empleados en el Comparador de Replicaciones

RESUMEN: En este apéndice se encuentran disponibles los códigos fuente usados en el procedimiento para comparar replicaciones.

C.1 Código fuente escrito en Java del programa *Transpose*

```
1  /**
2   * Programa Transpose <br>
3   * Transpose.java <br>
4   * <br>
5   * Propósito: Genera una serie de matrices traspuestas que se almacenan en
6   *   ficheros de
7   *   texto independientes. Como entrada este programa recibe un fichero de
8   *   texto con información
9   *   de los diferentes elementos de la situación experimental.
10  *
11  * @author Omar Salvador Gómez Gómez
12  * @version 1.0 12/21/2010
13  */
14
15 import java.io.File;
16 import java.util.List;
17 import java.util.ArrayList;
18 import java.util.Iterator;
19 import com.infomata.data.*;
20
21 public class Transpose{
22
23     public static void main(String[] argv) throws Exception{
24
25         DataFile read1 = DataFile.createReader("8859_1");
26         read1.setDataFormat(new CSVFormat());
```

```
25 read1.containsHeader(true);

27 int i1=1,i2=1,i3=1,i4=1,i5=1,i6=1,i7=1,i8=1,i9=1,i10=1,i11=1,i12=1,i13
    =1,i14=1,i15=1;
28 int i16=1,i17=1,i18=1,i19=1,i20=1,i21=1,i22=1,i23=1,i24=1,i25=1,i26=1,
    i27=1,i28=1,i29=1;
29 int i30=1,i31=1,i32=1,i33=1,i34=1,i35=1,i36=1,i37=1,i38=1;
30 String id="";
31 String partition=argv[0];
32 try {
33     read1.open(new File(partition));

35     for (DataRow row = read1.next(); row != null; row = read1.next()) {
36         id = row.getString("ID");
37         if (id.equals("1.1")) i1++;
38         else if (id.equals("1.2")) i2++;
39         else if (id.equals("1.3")) i3++;
40         else if (id.equals("1.4")) i4++;
41         else if (id.equals("1.5")) i5++;
42         else if (id.equals("2")) i6++;
43         else if (id.equals("3.1.1")) i7++;
44         else if (id.equals("3.2.1")) i8++;
45         else if (id.equals("3.2.2")) i9++;
46         else if (id.equals("3.2.3")) i10++;
47         else if (id.equals("3.2.4")) i11++;
48         else if (id.equals("3.2.5")) i12++;
49         else if (id.equals("3.2.6")) i13++;
50         else if (id.equals("3.3.1")) i14++;
51         else if (id.equals("3.4.1")) i15++;
52         else if (id.equals("3.4.2")) i16++;
53         else if (id.equals("3.5.1")) i17++;
54         else if (id.equals("3.5.2")) i18++;
55         else if (id.equals("3.5.3")) i19++;
56         else if (id.equals("4.1.1.1")) i20++;
57         else if (id.equals("4.1.1.2")) i21++;
58         else if (id.equals("4.1.1.3")) i22++;
59         else if (id.equals("4.1.1.4")) i23++;
60         else if (id.equals("4.1.1.5")) i24++;
61         else if (id.equals("4.1.2.1")) i25++;
62         else if (id.equals("4.1.2.2")) i26++;
63         else if (id.equals("4.1.2.3")) i27++;
64         else if (id.equals("4.1.2.4")) i28++;
65         else if (id.equals("4.1.2.5")) i29++;
66         else if (id.equals("4.2.1")) i30++;
67         else if (id.equals("4.2.2")) i31++;
68         else if (id.equals("5.1.1")) i32++;
69         else if (id.equals("5.1.2")) i33++;
70         else if (id.equals("5.1.3")) i34++;
```

```

71     else if (id.equals("5.2.1")) i35++;
72     else if (id.equals("5.2.2")) i36++;
73     else if (id.equals("5.2.3")) i37++;
74     else if (id.equals("5.2.4")) i38++;
75     }
76 }
77 finally {
78     read1.close();
79 }

81 DataFile read = DataFile.createReader("8859_1");
82 read.setDataFormat(new CSVFormat());
83 read.containsHeader(true);

85 try {
86     read.open(new File(partition));

88     List ids = new ArrayList();
89     List header = read.getHeaderList();
90     Iterator iterator = header.iterator();
91     int j = header.size();
92     String [][] m1_1 = new String[j][i1];
93     String [][] m1_2 = new String[j][i2];
94     String [][] m1_3 = new String[j][i3];
95     String [][] m1_4 = new String[j][i4];
96     String [][] m1_5 = new String[j][i5];
97     String [][] m2 = new String[j][i6];
98     String [][] m3_1_1 = new String[j][i7];
99     String [][] m3_2_1 = new String[j][i8];
100    String [][] m3_2_2 = new String[j][i9];
101    String [][] m3_2_3 = new String[j][i10];
102    String [][] m3_2_4 = new String[j][i11];
103    String [][] m3_2_5 = new String[j][i12];
104    String [][] m3_2_6 = new String[j][i13];
105    String [][] m3_3_1 = new String[j][i14];
106    String [][] m3_4_1 = new String[j][i15];
107    String [][] m3_4_2 = new String[j][i16];
108    String [][] m3_5_1 = new String[j][i17];
109    String [][] m3_5_2 = new String[j][i18];
110    String [][] m3_5_3 = new String[j][i19];
111    String [][] m4_1_1_1 = new String[j][i20];
112    String [][] m4_1_1_2 = new String[j][i21];
113    String [][] m4_1_1_3 = new String[j][i22];
114    String [][] m4_1_1_4 = new String[j][i23];
115    String [][] m4_1_1_5 = new String[j][i24];
116    String [][] m4_1_2_1 = new String[j][i25];
117    String [][] m4_1_2_2 = new String[j][i26];
118    String [][] m4_1_2_3 = new String[j][i27];

```

```
119 String [][] m4_1_2_4 = new String[j][i28];
120 String [][] m4_1_2_5 = new String[j][i29];
121 String [][] m4_2_1 = new String[j][i30];
122 String [][] m4_2_2 = new String[j][i31];
123 String [][] m5_1_1 = new String[j][i32];
124 String [][] m5_1_2 = new String[j][i33];
125 String [][] m5_1_3 = new String[j][i34];
126 String [][] m5_2_1 = new String[j][i35];
127 String [][] m5_2_2 = new String[j][i36];
128 String [][] m5_2_3 = new String[j][i37];
129 String [][] m5_2_4 = new String[j][i38];

131 int inc=0;
132 String s;
133 while ( iterator.hasNext() ){
134     s = (String)iterator.next();
135     m1_1[inc][0] = s;
136     m1_2[inc][0] = s;
137     m1_3[inc][0] = s;
138     m1_4[inc][0] = s;
139     m1_5[inc][0] = s;
140     m2[inc][0] = s;
141     m3_1_1[inc][0] = s;
142     m3_2_1[inc][0] = s;
143     m3_2_2[inc][0] = s;
144     m3_2_3[inc][0] = s;
145     m3_2_4[inc][0] = s;
146     m3_2_5[inc][0] = s;
147     m3_2_6[inc][0] = s;
148     m3_3_1[inc][0] = s;
149     m3_4_1[inc][0] = s;
150     m3_4_2[inc][0] = s;
151     m3_5_1[inc][0] = s;
152     m3_5_2[inc][0] = s;
153     m3_5_3[inc][0] = s;
154     m4_1_1_1[inc][0] = s;
155     m4_1_1_2[inc][0] = s;
156     m4_1_1_3[inc][0] = s;
157     m4_1_1_4[inc][0] = s;
158     m4_1_1_5[inc][0] = s;
159     m4_1_2_1[inc][0] = s;
160     m4_1_2_2[inc][0] = s;
161     m4_1_2_3[inc][0] = s;
162     m4_1_2_4[inc][0] = s;
163     m4_1_2_5[inc][0] = s;
164     m4_2_1[inc][0] = s;
165     m4_2_2[inc][0] = s;
166     m5_1_1[inc][0] = s;
```

```

167     m5_1_2[inc][0] = s;
168     m5_1_3[inc][0] = s;
169     m5_2_1[inc][0] = s;
170     m5_2_2[inc][0] = s;
171     m5_2_3[inc][0] = s;
172     m5_2_4[inc][0] = s;
173     inc++;

175 }

177 i1=1;i2=1;i3=1;i4=1;i5=1;i6=1;i7=1;i8=1;i9=1;i10=1;i11=1;i12=1;i13=1;i14
    =1;i15=1;
178 i16=1;i17=1;i18=1;i19=1;i20=1;i21=1;i22=1;i23=1;i24=1;i25=1;i26=1;i27=1;
    i28=1;i29=1;
179 i30=1;i31=1;i32=1;i33=1;i34=1;i35=1;i36=1;i37=1;i38=1;
180 id="";
181     for (DataRow row = read.next(); row != null; row = read.next()) {
182         id = row.getString("ID");
183         int x=0;
184         if (id.equals("1.1")){
185             for(x=0; x<j; x++){
186                 m1_1[x][i1] = row.getString(x);
187             }
188             i1++;
189             ids.add(id);
190         } else if (id.equals("1.2")) {
191             for(x=0; x<j; x++){
192                 m1_2[x][i2] = row.getString(x);
193             }
194             i2++;
195             ids.add(id);
196         } else if (id.equals("1.3")) {
197             for(x=0; x<j; x++){
198                 m1_3[x][i3] = row.getString(x);
199             }
200             i3++;
201             ids.add(id);
202         } else if (id.equals("1.4")) {
203             for(x=0; x<j; x++){
204                 m1_4[x][i4] = row.getString(x);
205             }
206             i4++;
207             ids.add(id);
208         } else if (id.equals("1.5")) {
209             for(x=0; x<j; x++){
210                 m1_5[x][i5] = row.getString(x);
211             }
212             i5++;

```



```
213     ids.add(id);
214 } else if (id.equals("2")) {
215     for(x=0; x<j; x++){
216         m2[x][i6] = row.getString(x);
217     }
218     i6++;
219     ids.add(id);
220 } else if (id.equals("3.1.1")) {
221     for(x=0; x<j; x++){
222         m3_1_1[x][i7] = row.getString(x);
223     }
224     i7++;
225     ids.add(id);
226 } else if (id.equals("3.2.1")) {
227     for(x=0; x<j; x++){
228         m3_2_1[x][i8] = row.getString(x);
229     }
230     i8++;
231     ids.add(id);
232 } else if (id.equals("3.2.2")) {
233     for(x=0; x<j; x++){
234         m3_2_2[x][i9] = row.getString(x);
235     }
236     i9++;
237     ids.add(id);
238 } else if (id.equals("3.2.3")) {
239     for(x=0; x<j; x++){
240         m3_2_3[x][i10] = row.getString(x);
241     }
242     i10++;
243     ids.add(id);
244 } else if (id.equals("3.2.4")) {
245     for(x=0; x<j; x++){
246         m3_2_4[x][i11] = row.getString(x);
247     }
248     i11++;
249     ids.add(id);
250 } else if (id.equals("3.2.5")) {
251     for(x=0; x<j; x++){
252         m3_2_5[x][i12] = row.getString(x);
253     }
254     i12++;
255     ids.add(id);
256 } else if (id.equals("3.2.6")) {
257     for(x=0; x<j; x++){
258         m3_2_6[x][i13] = row.getString(x);
259     }
260     i13++;
```

```
261     ids.add(id);
262 } else if (id.equals("3.3.1")) {
263     for(x=0; x<j; x++){
264         m3_3_1[x][i14] = row.getString(x);
265     }
266     i14++;
267     ids.add(id);
268 } else if (id.equals("3.4.1")) {
269     for(x=0; x<j; x++){
270         m3_4_1[x][i15] = row.getString(x);
271     }
272     i15++;
273     ids.add(id);
274 } else if (id.equals("3.4.2")) {
275     for(x=0; x<j; x++){
276         m3_4_2[x][i16] = row.getString(x);
277     }
278     i16++;
279     ids.add(id);
280 } else if (id.equals("3.5.1")) {
281     for(x=0; x<j; x++){
282         m3_5_1[x][i17] = row.getString(x);
283     }
284     i17++;
285     ids.add(id);
286 } else if (id.equals("3.5.2")) {
287     for(x=0; x<j; x++){
288         m3_5_2[x][i18] = row.getString(x);
289     }
290     i18++;
291     ids.add(id);
292 } else if (id.equals("3.5.3")) {
293     for(x=0; x<j; x++){
294         m3_5_3[x][i19] = row.getString(x);
295     }
296     i19++;
297     ids.add(id);
298 } else if (id.equals("4.1.1.1")) {
299     for(x=0; x<j; x++){
300         m4_1_1_1[x][i20] = row.getString(x);
301     }
302     i20++;
303     ids.add(id);
304 } else if (id.equals("4.1.1.2")) {
305     for(x=0; x<j; x++){
306         m4_1_1_2[x][i21] = row.getString(x);
307     }
308     i21++;
```

```
309     ids.add(id);
310 } else if (id.equals("4.1.1.3")) {
311     for(x=0; x<j; x++){
312         m4_1_1_3[x][i22] = row.getString(x);
313     }
314     i22++;
315     ids.add(id);
316 } else if (id.equals("4.1.1.4")) {
317     for(x=0; x<j; x++){
318         m4_1_1_4[x][i23] = row.getString(x);
319     }
320     i23++;
321     ids.add(id);
322 } else if (id.equals("4.1.1.5")) {
323     for(x=0; x<j; x++){
324         m4_1_1_5[x][i24] = row.getString(x);
325     }
326     i24++;
327     ids.add(id);
328 } else if (id.equals("4.1.2.1")) {
329     for(x=0; x<j; x++){
330         m4_1_2_1[x][i25] = row.getString(x);
331     }
332     i25++;
333     ids.add(id);
334 } else if (id.equals("4.1.2.2")) {
335     for(x=0; x<j; x++){
336         m4_1_2_2[x][i26] = row.getString(x);
337     }
338     i26++;
339     ids.add(id);
340 } else if (id.equals("4.1.2.3")) {
341     for(x=0; x<j; x++){
342         m4_1_2_3[x][i27] = row.getString(x);
343     }
344     i27++;
345     ids.add(id);
346 } else if (id.equals("4.1.2.4")) {
347     for(x=0; x<j; x++){
348         m4_1_2_4[x][i28] = row.getString(x);
349     }
350     i28++;
351     ids.add(id);
352 } else if (id.equals("4.1.2.5")) {
353     for(x=0; x<j; x++){
354         m4_1_2_5[x][i29] = row.getString(x);
355     }
356     i29++;
```

```

357     ids.add(id);
358 } else if (id.equals("4.2.1")) {
359     for(x=0; x<j; x++){
360         m4_2_1[x][i30] = row.getString(x);
361     }
362     i30++;
363     ids.add(id);
364 } else if (id.equals("4.2.2")) {
365     for(x=0; x<j; x++){
366         m4_2_2[x][i31] = row.getString(x);
367     }
368     i31++;
369     ids.add(id);
370 } else if (id.equals("5.1.1")) {
371     for(x=0; x<j; x++){
372         m5_1_1[x][i32] = row.getString(x);
373     }
374     i32++;
375     ids.add(id);
376 } else if (id.equals("5.1.2")) {
377     for(x=0; x<j; x++){
378         m5_1_2[x][i33] = row.getString(x);
379     }
380     i33++;
381     ids.add(id);
382 } else if (id.equals("5.1.3")) {
383     for(x=0; x<j; x++){
384         m5_1_3[x][i34] = row.getString(x);
385     }
386     i34++;
387     ids.add(id);
388 } else if (id.equals("5.2.1")) {
389     for(x=0; x<j; x++){
390         m5_2_1[x][i35] = row.getString(x);
391     }
392     i35++;
393     ids.add(id);
394 } else if (id.equals("5.2.2")) {
395     for(x=0; x<j; x++){
396         m5_2_2[x][i36] = row.getString(x);
397     }
398     i36++;
399     ids.add(id);
400 } else if (id.equals("5.2.3")) {
401     for(x=0; x<j; x++){
402         m5_2_3[x][i37] = row.getString(x);
403     }
404     i37++;

```

```
405     ids.add(id);
406 } else if (id.equals("5.2.4")) {
407     for(x=0; x<j; x++){
408         m5_2_4[x][i38] = row.getString(x);
409     }
410     i38++;
411     ids.add(id);
412 }
413 }

416 iterator = ids.iterator();
417 while ( iterator.hasNext() ){
418     id = (String)iterator.next();
419     if (id.equals("1.1")){
420         DataFile write = DataFile.createWriter("8859_1", false);
421         write.setDataFormat(new TabFormat());
422         try{
423             write.open(new File(id.concat(".txt")));
424             for (int x=0; x < m1_1.length; x++){
425                 DataRow row = write.next();
426                 for (int y=0; y < m1_1[x].length; y++){
427                     if (x==0 && y==0) row.add("");
428                     else if ( (!m1_1[x][y].equals("X")) && (!m1_1[x][y].equals("")) )
429                         row.add( m1_1[x][y] );
430                     else if (m1_1[x][y].equals("X")) row.add("1");
431                     else row.add("0");
432                 }
433             }
434         }
435         finally { write.close(); }
436     } else if (id.equals("1.2")) {
437         DataFile write = DataFile.createWriter("8859_1", false);
438         write.setDataFormat(new TabFormat());
439         try{
440             write.open(new File(id.concat(".txt")));
441             for (int x=0; x < m1_2.length; x++){
442                 DataRow row = write.next();
443                 for (int y=0; y < m1_2[x].length; y++){
444                     if (x==0 && y==0) row.add("");
445                     else if ( (!m1_2[x][y].equals("X")) && (!m1_2[x][y].equals("")) )
446                         row.add( m1_2[x][y] );
447                     else if (m1_2[x][y].equals("X")) row.add("1");
448                     else row.add("0");
449                 }
450             }
451         }
452         finally { write.close(); }
```

```

453 } else if (id.equals("1.3")) {
454     DataFile write = DataFile.createWriter("8859_1", false);
455     write.setDataFormat(new TabFormat());
456     try{
457         write.open(new File(id.concat(".txt")));
458         for (int x=0; x < m1_3.length; x++){
459             DataRow row = write.next();
460             for (int y=0; y < m1_3[x].length; y++){
461                 if (x==0 && y==0) row.add("");
462                 else if ( (!m1_3[x][y].equals("X")) && (!m1_3[x][y].equals(""))))
463                     row.add( m1_3[x][y] );
464                 else if (m1_3[x][y].equals("X")) row.add("1");
465                 else row.add("0");
466             }
467         }
468     }
469     finally { write.close(); }
470 } else if (id.equals("1.4")) {
471     DataFile write = DataFile.createWriter("8859_1", false);
472     write.setDataFormat(new TabFormat());
473     try{
474         write.open(new File(id.concat(".txt")));
475         for (int x=0; x < m1_4.length; x++){
476             DataRow row = write.next();
477             for (int y=0; y < m1_4[x].length; y++){
478                 if (x==0 && y==0) row.add("");
479                 else if ( (!m1_4[x][y].equals("X")) && (!m1_4[x][y].equals(""))))
480                     row.add( m1_4[x][y] );
481                 else if (m1_4[x][y].equals("X")) row.add("1");
482                 else row.add("0");
483             }
484         }
485     }
486     finally { write.close(); }
487 } else if (id.equals("1.5")) {
488     DataFile write = DataFile.createWriter("8859_1", false);
489     write.setDataFormat(new TabFormat());
490     try{
491         write.open(new File(id.concat(".txt")));
492         for (int x=0; x < m1_5.length; x++){
493             DataRow row = write.next();
494             for (int y=0; y < m1_5[x].length; y++){
495                 if (x==0 && y==0) row.add("");
496                 else if ( (!m1_5[x][y].equals("X")) && (!m1_5[x][y].equals(""))))
497                     row.add( m1_5[x][y] );
498                 else if (m1_5[x][y].equals("X")) row.add("1");
499                 else row.add("0");
500             }

```

```
501     }
502     }
503     finally { write.close(); }
504 } else if (id.equals("2")) {
505     DataFile write = DataFile.createWriter("8859_1", false);
506     write.setDataFormat(new TabFormat());
507     try{
508         write.open(new File(id.concat(".txt")));
509         for (int x=0; x < m2.length; x++){
510             DataRow row = write.next();
511             for (int y=0; y < m2[x].length; y++){
512                 if (x==0 && y==0) row.add("");
513                 else if ( (!m2[x][y].equals("X")) && (!m2[x][y].equals("")) )
514                     row.add( m2[x][y] );
515                 else if (m2[x][y].equals("X")) row.add("1");
516                 else row.add("0");
517             }
518         }
519     }
520     finally { write.close(); }
521 } else if (id.equals("3.1.1")) {
522     DataFile write = DataFile.createWriter("8859_1", false);
523     write.setDataFormat(new TabFormat());
524     try{
525         write.open(new File(id.concat(".txt")));
526         for (int x=0; x < m3_1_1.length; x++){
527             DataRow row = write.next();
528             for (int y=0; y < m3_1_1[x].length; y++){
529                 if (x==0 && y==0) row.add("");
530                 else if ( (!m3_1_1[x][y].equals("X")) && (!m3_1_1[x][y].equals(""))
531                     ))
532                     row.add( m3_1_1[x][y] );
533                 else if (m3_1_1[x][y].equals("X")) row.add("1");
534                 else row.add("0");
535             }
536         }
537     }
538     finally { write.close(); }
539 } else if (id.equals("3.2.1")) {
540     DataFile write = DataFile.createWriter("8859_1", false);
541     write.setDataFormat(new TabFormat());
542     try{
543         write.open(new File(id.concat(".txt")));
544         for (int x=0; x < m3_2_1.length; x++){
545             DataRow row = write.next();
546             for (int y=0; y < m3_2_1[x].length; y++){
547                 if (x==0 && y==0) row.add("");
```

```

547         else if ( (!m3_2_1[x][y].equals("X")) && (!m3_2_1[x][y].equals(""))
548             ))
549             row.add( m3_2_1[x][y] );
550         else if (m3_2_1[x][y].equals("X")) row.add("1");
551         else row.add("0");
552     }
553 }
554 finally { write.close(); }
555 } else if (id.equals("3.2.2")) {
556     DataFile write = DataFile.createWriter("8859_1", false);
557     write.setDataFormat(new TabFormat());
558     try{
559         write.open(new File(id.concat(".txt")));
560         for (int x=0; x < m3_2_2.length; x++){
561             DataRow row = write.next();
562             for (int y=0; y < m3_2_2[x].length; y++){
563                 if (x==0 && y==0) row.add("");
564                 else if ( (!m3_2_2[x][y].equals("X")) && (!m3_2_2[x][y].equals(""))
565                     ))
566                     row.add( m3_2_2[x][y] );
567                 else if (m3_2_2[x][y].equals("X")) row.add("1");
568                 else row.add("0");
569             }
570         }
571         finally { write.close(); }
572     } else if (id.equals("3.2.3")) {
573         DataFile write = DataFile.createWriter("8859_1", false);
574         write.setDataFormat(new TabFormat());
575         try{
576             write.open(new File(id.concat(".txt")));
577             for (int x=0; x < m3_2_3.length; x++){
578                 DataRow row = write.next();
579                 for (int y=0; y < m3_2_3[x].length; y++){
580                     if (x==0 && y==0) row.add("");
581                     else if ( (!m3_2_3[x][y].equals("X")) && (!m3_2_3[x][y].equals(""))
582                         ))
583                         row.add( m3_2_3[x][y] );
584                     else if (m3_2_3[x][y].equals("X")) row.add("1");
585                     else row.add("0");
586                 }
587             }
588             finally { write.close(); }
589         } else if (id.equals("3.2.4")) {
590             DataFile write = DataFile.createWriter("8859_1", false);
591             write.setDataFormat(new TabFormat());

```



```
592     try{
593         write.open(new File(id.concat(".txt")));
594         for (int x=0; x < m3_2_4.length; x++){
595             DataRow row = write.next();
596             for (int y=0; y < m3_2_4[x].length; y++){
597                 if (x==0 && y==0) row.add("");
598                 else if ( (!m3_2_4[x][y].equals("X")) && (!m3_2_4[x][y].equals(""))
599                     ))
600                     row.add( m3_2_4[x][y] );
601                 else if (m3_2_4[x][y].equals("X")) row.add("1");
602                 else row.add("0");
603             }
604         }
605         finally { write.close(); }
606     } else if (id.equals("3.2.5")) {
607         DataFile write = DataFile.createWriter("8859_1", false);
608         write.setDataFormat(new TabFormat());
609         try{
610             write.open(new File(id.concat(".txt")));
611             for (int x=0; x < m3_2_5.length; x++){
612                 DataRow row = write.next();
613                 for (int y=0; y < m3_2_5[x].length; y++){
614                     if (x==0 && y==0) row.add("");
615                     else if ( (!m3_2_5[x][y].equals("X")) && (!m3_2_5[x][y].equals(""))
616                         ))
617                         row.add( m3_2_5[x][y] );
618                     else if (m3_2_5[x][y].equals("X")) row.add("1");
619                     else row.add("0");
620                 }
621             }
622             finally { write.close(); }
623     } else if (id.equals("3.2.6")) {
624         DataFile write = DataFile.createWriter("8859_1", false);
625         write.setDataFormat(new TabFormat());
626         try{
627             write.open(new File(id.concat(".txt")));
628             for (int x=0; x < m3_2_6.length; x++){
629                 DataRow row = write.next();
630                 for (int y=0; y < m3_2_6[x].length; y++){
631                     if (x==0 && y==0) row.add("");
632                     else if ( (!m3_2_6[x][y].equals("X")) && (!m3_2_6[x][y].equals(""))
633                         ))
634                         row.add( m3_2_6[x][y] );
635                     else if (m3_2_6[x][y].equals("X")) row.add("1");
636                     else row.add("0");
637                 }
638             }
639             finally { write.close(); }
640         }
641     }
```

```

637     }
638 }
639 finally { write.close(); }
640 } else if (id.equals("3.3.1")) {
641     DataFile write = DataFile.createWriter("8859_1", false);
642     write.setDataFormat(new TabFormat());
643     try{
644         write.open(new File(id.concat(".txt")));
645         for (int x=0; x < m3_3_1.length; x++){
646             DataRow row = write.next();
647             for (int y=0; y < m3_3_1[x].length; y++){
648                 if (x==0 && y==0) row.add("");
649                 else if ( (!m3_3_1[x][y].equals("X")) && (!m3_3_1[x][y].equals(""))
650                     ))
651                     row.add( m3_3_1[x][y] );
652                 else if (m3_3_1[x][y].equals("X")) row.add("1");
653                 else row.add("0");
654             }
655         }
656         finally { write.close(); }
657     } else if (id.equals("3.4.1")) {
658         DataFile write = DataFile.createWriter("8859_1", false);
659         write.setDataFormat(new TabFormat());
660         try{
661             write.open(new File(id.concat(".txt")));
662             for (int x=0; x < m3_4_1.length; x++){
663                 DataRow row = write.next();
664                 for (int y=0; y < m3_4_1[x].length; y++){
665                     if (x==0 && y==0) row.add("");
666                     else if ( (!m3_4_1[x][y].equals("X")) && (!m3_4_1[x][y].equals(""))
667                         ))
668                         row.add( m3_4_1[x][y] );
669                     else if (m3_4_1[x][y].equals("X")) row.add("1");
670                     else row.add("0");
671                 }
672             }
673             finally { write.close(); }
674     } else if (id.equals("3.4.2")) {
675         DataFile write = DataFile.createWriter("8859_1", false);
676         write.setDataFormat(new TabFormat());
677         try{
678             write.open(new File(id.concat(".txt")));
679             for (int x=0; x < m3_4_2.length; x++){
680                 DataRow row = write.next();
681                 for (int y=0; y < m3_4_2[x].length; y++){
682                     if (x==0 && y==0) row.add("");

```

```
683         else if ( (!m3_4_2[x][y].equals("X")) && (!m3_4_2[x][y].equals(""))
684             ))
685             row.add( m3_4_2[x][y] );
686         else if (m3_4_2[x][y].equals("X")) row.add("1");
687         else row.add("0");
688     }
689 }
690 finally { write.close(); }
691 } else if (id.equals("3.5.1")) {
692     DataFile write = DataFile.createWriter("8859_1", false);
693     write.setDataFormat(new TabFormat());
694     try{
695         write.open(new File(id.concat(".txt")));
696         for (int x=0; x < m3_5_1.length; x++){
697             DataRow row = write.next();
698             for (int y=0; y < m3_5_1[x].length; y++){
699                 if (x==0 && y==0) row.add("");
700                 else if ( (!m3_5_1[x][y].equals("X")) && (!m3_5_1[x][y].equals(""))
701                     ))
702                     row.add( m3_5_1[x][y] );
703                 else if (m3_5_1[x][y].equals("X")) row.add("1");
704                 else row.add("0");
705             }
706         }
707         finally { write.close(); }
708     } else if (id.equals("3.5.2")) {
709         DataFile write = DataFile.createWriter("8859_1", false);
710         write.setDataFormat(new TabFormat());
711         try{
712             write.open(new File(id.concat(".txt")));
713             for (int x=0; x < m3_5_2.length; x++){
714                 DataRow row = write.next();
715                 for (int y=0; y < m3_5_2[x].length; y++){
716                     if (x==0 && y==0) row.add("");
717                     else if ( (!m3_5_2[x][y].equals("X")) && (!m3_5_2[x][y].equals(""))
718                         ))
719                         row.add( m3_5_2[x][y] );
720                     else if (m3_5_2[x][y].equals("X")) row.add("1");
721                     else row.add("0");
722                 }
723             }
724             finally { write.close(); }
725         } else if (id.equals("3.5.3")) {
726             DataFile write = DataFile.createWriter("8859_1", false);
727             write.setDataFormat(new TabFormat());
```

```

728     try{
729         write.open(new File(id.concat(".txt")));
730         for (int x=0; x < m3_5_3.length; x++){
731             DataRow row = write.next();
732             for (int y=0; y < m3_5_3[x].length; y++){
733                 if (x==0 && y==0) row.add("");
734                 else if ( (!m3_5_3[x][y].equals("X")) && (!m3_5_3[x][y].equals(""))
735                     ))
736                     row.add( m3_5_3[x][y] );
737                 else if (m3_5_3[x][y].equals("X")) row.add("1");
738                 else row.add("0");
739             }
740         }
741         finally { write.close(); }
742     } else if (id.equals("4.1.1.1")) {
743         DataFile write = DataFile.createWriter("8859_1", false);
744         write.setDataFormat(new TabFormat());
745         try{
746             write.open(new File(id.concat(".txt")));
747             for (int x=0; x < m4_1_1_1.length; x++){
748                 DataRow row = write.next();
749                 for (int y=0; y < m4_1_1_1[x].length; y++){
750                     if (x==0 && y==0) row.add("");
751                     else if ( (!m4_1_1_1[x][y].equals("X")) && (!m4_1_1_1[x][y].equals(
752                         "")))
753                         row.add( m4_1_1_1[x][y] );
754                     else if (m4_1_1_1[x][y].equals("X")) row.add("1");
755                     else row.add("0");
756                 }
757             }
758             finally { write.close(); }
759     } else if (id.equals("4.1.1.2")) {
760         DataFile write = DataFile.createWriter("8859_1", false);
761         write.setDataFormat(new TabFormat());
762         try{
763             write.open(new File(id.concat(".txt")));
764             for (int x=0; x < m4_1_1_2.length; x++){
765                 DataRow row = write.next();
766                 for (int y=0; y < m4_1_1_2[x].length; y++){
767                     if (x==0 && y==0) row.add("");
768                     else if ( (!m4_1_1_2[x][y].equals("X")) && (!m4_1_1_2[x][y].equals(
769                         "")))
770                         row.add( m4_1_1_2[x][y] );
771                     else if (m4_1_1_2[x][y].equals("X")) row.add("1");
772                     else row.add("0");
773                 }

```

```
773     }
774   }
775   finally { write.close(); }
776 } else if (id.equals("4.1.1.3")) {
777   DataFile write = DataFile.createWriter("8859_1", false);
778   write.setDataFormat(new TabFormat());
779   try{
780     write.open(new File(id.concat(".txt")));
781     for (int x=0; x < m4_1_1_3.length; x++){
782       DataRow row = write.next();
783       for (int y=0; y < m4_1_1_3[x].length; y++){
784         if (x==0 && y==0) row.add("");
785         else if ( (!m4_1_1_3[x][y].equals("X")) && (!m4_1_1_3[x][y].equals
786           ("")))
787           row.add( m4_1_1_3[x][y] );
788         else if (m4_1_1_3[x][y].equals("X")) row.add("1");
789         else row.add("0");
790       }
791     }
792     finally { write.close(); }
793 } else if (id.equals("4.1.1.4")) {
794   DataFile write = DataFile.createWriter("8859_1", false);
795   write.setDataFormat(new TabFormat());
796   try{
797     write.open(new File(id.concat(".txt")));
798     for (int x=0; x < m4_1_1_4.length; x++){
799       DataRow row = write.next();
800       for (int y=0; y < m4_1_1_4[x].length; y++){
801         if (x==0 && y==0) row.add("");
802         else if ( (!m4_1_1_4[x][y].equals("X")) && (!m4_1_1_4[x][y].equals
803           ("")))
804           row.add( m4_1_1_4[x][y] );
805         else if (m4_1_1_4[x][y].equals("X")) row.add("1");
806         else row.add("0");
807       }
808     }
809     finally { write.close(); }
810 } else if (id.equals("4.1.1.5")) {
811   DataFile write = DataFile.createWriter("8859_1", false);
812   write.setDataFormat(new TabFormat());
813   try{
814     write.open(new File(id.concat(".txt")));
815     for (int x=0; x < m4_1_1_5.length; x++){
816       DataRow row = write.next();
817       for (int y=0; y < m4_1_1_5[x].length; y++){
818         if (x==0 && y==0) row.add("");
```

```

819         else if ( (!m4_1_1_5[x][y].equals("X")) && (!m4_1_1_5[x][y].equals
820             ("")))
821             row.add( m4_1_1_5[x][y] );
822         else if (m4_1_1_5[x][y].equals("X")) row.add("1");
823         else row.add("0");
824     }
825 }
826 finally { write.close(); }
827 } else if (id.equals("4.1.2.1")) {
828     DataFile write = DataFile.createWriter("8859_1", false);
829     write.setDataFormat(new TabFormat());
830     try{
831         write.open(new File(id.concat(".txt")));
832         for (int x=0; x < m4_1_2_1.length; x++){
833             DataRow row = write.next();
834             for (int y=0; y < m4_1_2_1[x].length; y++){
835                 if (x==0 && y==0) row.add("");
836                 else if ( (!m4_1_2_1[x][y].equals("X")) && (!m4_1_2_1[x][y].equals
837                     ("")))
838                     row.add( m4_1_2_1[x][y] );
839                 else if (m4_1_2_1[x][y].equals("X")) row.add("1");
840                 else row.add("0");
841             }
842         }
843         finally { write.close(); }
844     } else if (id.equals("4.1.2.2")) {
845         DataFile write = DataFile.createWriter("8859_1", false);
846         write.setDataFormat(new TabFormat());
847         try{
848             write.open(new File(id.concat(".txt")));
849             for (int x=0; x < m4_1_2_2.length; x++){
850                 DataRow row = write.next();
851                 for (int y=0; y < m4_1_2_2[x].length; y++){
852                     if (x==0 && y==0) row.add("");
853                     else if ( (!m4_1_2_2[x][y].equals("X")) && (!m4_1_2_2[x][y].equals
854                         ("")))
855                         row.add( m4_1_2_2[x][y] );
856                     else if (m4_1_2_2[x][y].equals("X")) row.add("1");
857                     else row.add("0");
858                 }
859             }
860             finally { write.close(); }
861         } else if (id.equals("4.1.2.3")) {
862             DataFile write = DataFile.createWriter("8859_1", false);
863             write.setDataFormat(new TabFormat());

```

```
864     try{
865         write.open(new File(id.concat(".txt")));
866         for (int x=0; x < m4_1_2_3.length; x++){
867             DataRow row = write.next();
868             for (int y=0; y < m4_1_2_3[x].length; y++){
869                 if (x==0 && y==0) row.add("");
870                 else if ( (!m4_1_2_3[x][y].equals("X")) && (!m4_1_2_3[x][y].equals
871                     ("")))
872                     row.add( m4_1_2_3[x][y] );
873                 else if (m4_1_2_3[x][y].equals("X")) row.add("1");
874                 else row.add("0");
875             }
876         }
877         finally { write.close(); }
878     } else if (id.equals("4.1.2.4")) {
879         DataFile write = DataFile.createWriter("8859_1", false);
880         write.setDataFormat(new TabFormat());
881         try{
882             write.open(new File(id.concat(".txt")));
883             for (int x=0; x < m4_1_2_4.length; x++){
884                 DataRow row = write.next();
885                 for (int y=0; y < m4_1_2_4[x].length; y++){
886                     if (x==0 && y==0) row.add("");
887                     else if ( (!m4_1_2_4[x][y].equals("X")) && (!m4_1_2_4[x][y].equals
888                         ("")))
889                         row.add( m4_1_2_4[x][y] );
890                     else if (m4_1_2_4[x][y].equals("X")) row.add("1");
891                     else row.add("0");
892                 }
893             }
894             finally { write.close(); }
895         } else if (id.equals("4.1.2.5")) {
896             DataFile write = DataFile.createWriter("8859_1", false);
897             write.setDataFormat(new TabFormat());
898             try{
899                 write.open(new File(id.concat(".txt")));
900                 for (int x=0; x < m4_1_2_5.length; x++){
901                     DataRow row = write.next();
902                     for (int y=0; y < m4_1_2_5[x].length; y++){
903                         if (x==0 && y==0) row.add("");
904                         else if ( (!m4_1_2_5[x][y].equals("X")) && (!m4_1_2_5[x][y].equals
905                             ("")))
906                             row.add( m4_1_2_5[x][y] );
907                         else if (m4_1_2_5[x][y].equals("X")) row.add("1");
908                         else row.add("0");
909                     }
910                 }
911             }
912             finally { write.close(); }
```

```

909     }
910 }
911 finally { write.close(); }
912 } else if (id.equals("4.2.1")) {
913     DataFile write = DataFile.createWriter("8859_1", false);
914     write.setDataFormat(new TabFormat());
915     try{
916         write.open(new File(id.concat(".txt")));
917         for (int x=0; x < m4_2_1.length; x++){
918             DataRow row = write.next();
919             for (int y=0; y < m4_2_1[x].length; y++){
920                 if (x==0 && y==0) row.add("");
921                 else if ( (!m4_2_1[x][y].equals("X")) && (!m4_2_1[x][y].equals(""))
922                     ))
923                     row.add( m4_2_1[x][y] );
924                 else if (m4_2_1[x][y].equals("X")) row.add("1");
925                 else row.add("0");
926             }
927         }
928         finally { write.close(); }
929     } else if (id.equals("4.2.2")) {
930         DataFile write = DataFile.createWriter("8859_1", false);
931         write.setDataFormat(new TabFormat());
932         try{
933             write.open(new File(id.concat(".txt")));
934             for (int x=0; x < m4_2_2.length; x++){
935                 DataRow row = write.next();
936                 for (int y=0; y < m4_2_2[x].length; y++){
937                     if (x==0 && y==0) row.add("");
938                     else if ( (!m4_2_2[x][y].equals("X")) && (!m4_2_2[x][y].equals(""))
939                         ))
940                         row.add( m4_2_2[x][y] );
941                     else if (m4_2_2[x][y].equals("X")) row.add("1");
942                     else row.add("0");
943                 }
944             }
945             finally { write.close(); }
946         } else if (id.equals("5.1.1")) {
947             DataFile write = DataFile.createWriter("8859_1", false);
948             write.setDataFormat(new TabFormat());
949             try{
950                 write.open(new File(id.concat(".txt")));
951                 for (int x=0; x < m5_1_1.length; x++){
952                     DataRow row = write.next();
953                     for (int y=0; y < m5_1_1[x].length; y++){
954                         if (x==0 && y==0) row.add("");

```



```
955         else if ( (!m5_1_1[x][y].equals("X")) && (!m5_1_1[x][y].equals(""))
956             ))
957             row.add( m5_1_1[x][y] );
958         else if (m5_1_1[x][y].equals("X")) row.add("1");
959         else row.add("0");
960     }
961 }
962 finally { write.close(); }
963 } else if (id.equals("5.1.2")) {
964     DataFile write = DataFile.createWriter("8859_1", false);
965     write.setDataFormat(new TabFormat());
966     try{
967         write.open(new File(id.concat(".txt")));
968         for (int x=0; x < m5_1_2.length; x++){
969             DataRow row = write.next();
970             for (int y=0; y < m5_1_2[x].length; y++){
971                 if (x==0 && y==0) row.add("");
972                 else if ( (!m5_1_2[x][y].equals("X")) && (!m5_1_2[x][y].equals(""))
973                     ))
974                     row.add( m5_1_2[x][y] );
975                 else if (m5_1_2[x][y].equals("X")) row.add("1");
976                 else row.add("0");
977             }
978         }
979         finally { write.close(); }
980     } else if (id.equals("5.1.3")) {
981         DataFile write = DataFile.createWriter("8859_1", false);
982         write.setDataFormat(new TabFormat());
983         try{
984             write.open(new File(id.concat(".txt")));
985             for (int x=0; x < m5_1_3.length; x++){
986                 DataRow row = write.next();
987                 for (int y=0; y < m5_1_3[x].length; y++){
988                     if (x==0 && y==0) row.add("");
989                     else if ( (!m5_1_3[x][y].equals("X")) && (!m5_1_3[x][y].equals(""))
990                         ))
991                         row.add( m5_1_3[x][y] );
992                     else if (m5_1_3[x][y].equals("X")) row.add("1");
993                     else row.add("0");
994                 }
995             }
996             finally { write.close(); }
997         } else if (id.equals("5.2.1")) {
998             DataFile write = DataFile.createWriter("8859_1", false);
999             write.setDataFormat(new TabFormat());
```

```

1000     try{
1001         write.open(new File(id.concat(".txt")));
1002         for (int x=0; x < m5_2_1.length; x++){
1003             DataRow row = write.next();
1004             for (int y=0; y < m5_2_1[x].length; y++){
1005                 if (x==0 && y==0) row.add("");
1006                 else if ( (!m5_2_1[x][y].equals("X")) && (!m5_2_1[x][y].equals(""))
1007                     ))
1008                     row.add( m5_2_1[x][y] );
1009                 else if (m5_2_1[x][y].equals("X")) row.add("1");
1010                 else row.add("0");
1011             }
1012         }
1013         finally { write.close(); }
1014     } else if (id.equals("5.2.2")) {
1015         DataFile write = DataFile.createWriter("8859_1", false);
1016         write.setDataFormat(new TabFormat());
1017         try{
1018             write.open(new File(id.concat(".txt")));
1019             for (int x=0; x < m5_2_2.length; x++){
1020                 DataRow row = write.next();
1021                 for (int y=0; y < m5_2_2[x].length; y++){
1022                     if (x==0 && y==0) row.add("");
1023                     else if ( (!m5_2_2[x][y].equals("X")) && (!m5_2_2[x][y].equals(""))
1024                         ))
1025                         row.add( m5_2_2[x][y] );
1026                     else if (m5_2_2[x][y].equals("X")) row.add("1");
1027                     else row.add("0");
1028                 }
1029             }
1030             finally { write.close(); }
1031     } else if (id.equals("5.2.3")) {
1032         DataFile write = DataFile.createWriter("8859_1", false);
1033         write.setDataFormat(new TabFormat());
1034         try{
1035             write.open(new File(id.concat(".txt")));
1036             for (int x=0; x < m5_2_3.length; x++){
1037                 DataRow row = write.next();
1038                 for (int y=0; y < m5_2_3[x].length; y++){
1039                     if (x==0 && y==0) row.add("");
1040                     else if ( (!m5_2_3[x][y].equals("X")) && (!m5_2_3[x][y].equals(""))
1041                         ))
1042                         row.add( m5_2_3[x][y] );
1043                     else if (m5_2_3[x][y].equals("X")) row.add("1");
1044                     else row.add("0");
1045                 }
1046             }
1047             finally { write.close(); }
1048         }
1049     }

```

```
1045     }
1046     }
1047     finally { write.close(); }
1048 } else if (id.equals("5.2.4")) {
1049     DataFile write = DataFile.createWriter("8859_1", false);
1050     write.setDataFormat(new TabFormat());
1051     try{
1052         write.open(new File(id.concat(".txt")));
1053         for (int x=0; x < m5_2_4.length; x++){
1054             DataRow row = write.next();
1055             for (int y=0; y < m5_2_4[x].length; y++){
1056                 if (x==0 && y==0) row.add("");
1057                 else if ( (!m5_2_4[x][y].equals("X")) && (!m5_2_4[x][y].equals(""))
1058                     ))
1059                     row.add( m5_2_4[x][y] );
1059                 else if (m5_2_4[x][y].equals("X")) row.add("1");
1060                 else row.add("0");
1061             }
1062         }
1063     }
1064     finally { write.close(); }
1065 }
1066 }
1067 }
1068 finally {
1069     read.close();
1070 }
1071 }
1072 }
```

Listado C.1: Transpose.java

C.2 Fichero *r_script01.txt* escrito en el lenguaje R

```

1  #-- Author: Omar S. Gómez G.
2  #-- Fecha: Mar.2011
3  #-- Versiion: 1.0

5  #----- Funciones para generar las representaciones visuales de las
    particiones

7  PrintDendrogram <- function(file, dendrogram, msg){
8      postscript(file, width = 6, height = 6,
9          horizontal = FALSE, onefile = FALSE, paper = "special",
10         family = "Helvetica")
11      plot(dendrogram,nodePar=list(pch = 2:1,cex=.4*2:1, col = 2:3), ylim=c
12          (0,1), main= msg)
13      dev.off()
14      return(TRUE)
15  }

16  PrintPCO <- function(file, dm){
17      postscript(file, width = 6, height = 6,
18          horizontal = FALSE, onefile = FALSE, paper = "special",
19          family = "Helvetica", pointsize=10)
20      #scatter(dm)
21      s.label(dm$li)
22      dev.off()
23      return(TRUE)
24  }

25  ##----- Fin de funciones
    -----

27  #-- Directorio donde se encuentran las tablas de contingencia y directorio
    de salida
28  dir = "part01/contingency_tables/"
29  eps = "part01/sim_comp/"

31  #-- Genera por cada (sub)elemento de la situación experimental, una matriz
    de distancias y
32  #-- almacenar en una variable identificada por cada (sub)elemento el
    contenido de la matriz
33  #-- de distancias.

35  s_1.1 = dist.binary(read.table(paste(dir,"1.1.txt",sep = "")),method=1,
    diag=TRUE)
36  s_1.2 = dist.binary(read.table(paste(dir,"1.2.txt",sep = "")),method=1,
    diag=TRUE)
37  s_1.3 = dist.binary(read.table(paste(dir,"1.3.txt",sep = "")),method=1,
    diag=TRUE)

```

```
38 s_1.4 = dist.binary(read.table(paste(dir,"1.4.txt",sep = "")),method=1,
    diag=TRUE)
39 s_1.5 = dist.binary(read.table(paste(dir,"1.5.txt",sep = "")),method=1,
    diag=TRUE)
40 s_2 = dist.binary(read.table(paste(dir,"2.txt",sep = "")),method=1,diag=
    TRUE)
41 s_3.1.1 = dist.binary(read.table(paste(dir,"3.1.1.txt",sep = "")),method
    =1,diag=TRUE)
42 s_3.2.1 = dist.binary(read.table(paste(dir,"3.2.1.txt",sep = "")),method
    =1,diag=TRUE)
43 s_3.2.2 = dist.binary(read.table(paste(dir,"3.2.2.txt",sep = "")),method
    =1,diag=TRUE)
44 s_3.2.3 = dist.binary(read.table(paste(dir,"3.2.3.txt",sep = "")),method
    =1,diag=TRUE)
45 s_3.2.4 = dist.binary(read.table(paste(dir,"3.2.4.txt",sep = "")),method
    =1,diag=TRUE)
46 s_3.2.5 = dist.binary(read.table(paste(dir,"3.2.5.txt",sep = "")),method
    =1,diag=TRUE)
47 s_3.2.6 = dist.binary(read.table(paste(dir,"3.2.6.txt",sep = "")),method
    =1,diag=TRUE)
48 s_3.3.1 = dist.binary(read.table(paste(dir,"3.3.1.txt",sep = "")),method
    =1,diag=TRUE)
49 s_3.4.1 = dist.binary(read.table(paste(dir,"3.4.1.txt",sep = "")),method
    =1,diag=TRUE)
50 s_3.4.2 = dist.binary(read.table(paste(dir,"3.4.2.txt",sep = "")),method
    =1,diag=TRUE)
51 s_3.5.1 = dist.binary(read.table(paste(dir,"3.5.1.txt",sep = "")),method
    =1,diag=TRUE)
52 s_3.5.2 = dist.binary(read.table(paste(dir,"3.5.2.txt",sep = "")),method
    =1,diag=TRUE)
53 s_3.5.3 = dist.binary(read.table(paste(dir,"3.5.3.txt",sep = "")),method
    =1,diag=TRUE)
54 s_4.1.1.1 = dist.binary(read.table(paste(dir,"4.1.1.1.txt",sep = "")),
    method=1,diag=TRUE)
55 s_4.1.1.2 = dist.binary(read.table(paste(dir,"4.1.1.2.txt",sep = "")),
    method=1,diag=TRUE)
56 s_4.1.1.3 = dist.binary(read.table(paste(dir,"4.1.1.3.txt",sep = "")),
    method=1,diag=TRUE)
57 s_4.1.1.4 = dist.binary(read.table(paste(dir,"4.1.1.4.txt",sep = "")),
    method=1,diag=TRUE)
58 s_4.1.1.5 = dist.binary(read.table(paste(dir,"4.1.1.5.txt",sep = "")),
    method=1,diag=TRUE)
59 s_4.1.2.1 = dist.binary(read.table(paste(dir,"4.1.2.1.txt",sep = "")),
    method=1,diag=TRUE)
60 s_4.1.2.2 = dist.binary(read.table(paste(dir,"4.1.2.2.txt",sep = "")),
    method=1,diag=TRUE)
61 s_4.1.2.3 = dist.binary(read.table(paste(dir,"4.1.2.3.txt",sep = "")),
    method=1,diag=TRUE)
```

```

62 s_4.1.2.4 = dist.binary(read.table(paste(dir,"4.1.2.4.txt",sep = "")),
    method=1,diag=TRUE)
63 s_4.1.2.5 = dist.binary(read.table(paste(dir,"4.1.2.5.txt",sep = "")),
    method=1,diag=TRUE)
64 s_4.2.1 = dist.binary(read.table(paste(dir,"4.2.1.txt",sep = "")),method
    =1,diag=TRUE)
65 s_5.1.1 = dist.binary(read.table(paste(dir,"5.1.1.txt",sep = "")),method
    =1,diag=TRUE)
66 s_5.1.2 = dist.binary(read.table(paste(dir,"5.1.2.txt",sep = "")),method
    =1,diag=TRUE)
67 s_5.1.3 = dist.binary(read.table(paste(dir,"5.1.3.txt",sep = "")),method
    =1,diag=TRUE)
68 s_5.2.1 = dist.binary(read.table(paste(dir,"5.2.1.txt",sep = "")),method
    =1,diag=TRUE)
69 s_5.2.2 = dist.binary(read.table(paste(dir,"5.2.2.txt",sep = "")),method
    =1,diag=TRUE)
70 s_5.2.3 = dist.binary(read.table(paste(dir,"5.2.3.txt",sep = "")),method
    =1,diag=TRUE)

72 #---Combinación de matrices de distancias
73 #---Combina matrices de forma jerárquica, primero a nivel de (sub)
    elementos.
74 experimentadores = (s_1.1+s_1.2+s_1.3+s_1.4+s_1.5)/5

76 sitio = s_2

78 protocolo = (s_3.1.1+(s_3.2.1+s_3.2.2+s_3.2.3+s_3.2.4+s_3.2.5+s_3.2.6)/6+
    s_3.3.1+
79 (s_3.4.1+s_3.4.2)/2+(s_3.5.1+s_3.5.2+s_3.5.3)/3)/5

81 operacionalizaciones = (((s_4.1.1.1+s_4.1.1.2+s_4.1.1.3+s_4.1.1.4+s_4
    .1.1.5)/5 +
82 (s_4.1.2.1+s_4.1.2.2+s_4.1.2.3+s_4.1.2.4+s_4.1.2.5 )/5)/2+s_4.2.1)
    /2
83 poblaciones = ( (s_5.1.1+s_5.1.2+s_5.1.3)/3+(s_5.2.1+s_5.2.2+s_5.2.3)/3)/2

86 #--- Combina matrices a nivel de secciones que componen la situación
    experimental
87 sit_exp = (protocolo+experimentadores+sitio+operacionalizaciones+
    poblaciones)/5

89 #--- Convertir matriz a matriz de distancias en un espacio euclídeo
90 ifelse(is.euclid(sit_exp), ec<-sit_exp, ec<-quasieuclid(sit_exp))
91 ec<-as.matrix(ec)
92 colnames(ec)<-labels(sit_exp)
93 rownames(ec)<-labels(sit_exp)
94 ec<-as.dist(ec,diag=TRUE)

```

```
96 #--- Genera el dendrograma
97 d_sit_exp = as.dendrogram(hclust(ec,method="average") )

99 #--- Realiza análisis escalado multidimensional métrico
100 sit_exp_pco <- dudi.pco(ec, scannf = FALSE, nf = 3)

102 #--- Genera representaciones visuales en formato postscript
103 PrintDendrogram(paste(eps,"_dendrogram_se_part01.eps",sep = ""),d_sit_exp,
  "")
104 PrintPCO(paste(eps,"_pco_se_part01.eps",sep = ""),sit_exp_pco)
```

Listado C.2: r_script01.txt

C.3 Fichero *r_script02.txt* escrito en el lenguaje R

```

1  #-- Author: Omar S. Gómez G.
2  #-- Fecha: Mar.2011
3  #-- Versiion: 1.0

5  PrintMeta <- function(file, meta){
6    #postscript(file, width = 4, height = 5,horizontal = FALSE, onefile =
7      FALSE, paper = "special",
8      postscript(file, width = 8.5, height = 3,horizontal = FALSE, onefile =
9        FALSE, paper = "special",
10         family = "Helvetica")

11    forest.meta(meta, comb.fixed=FALSE, comb.random=TRUE, fontsize=8,
12      plotwidth=unit(2.8,"inches"),
13      print.Q=TRUE, print.pval.Q=TRUE, just="left",xlim="s",leftcols=c("
14        studlab", "n.e",
15        "mean.e","sd.e","n.c","mean.c","sd.c"), lab.e.attach.to.col="sd.e",
16        lab.c.attach.to.col="sd.c",digits=4)

17    #plot.meta(meta, comb.fixed=FALSE, comb.random=TRUE, text.random="", xlab
18      = "")

19    dev.off()
20    return(TRUE)
21  }

22  PrintCC <- function(file, mean, studies, title, efect){
23    postscript(file, width = 5.5, height = 5, horizontal = FALSE, onefile =
24      FALSE, paper = "special",
25      family = "Helvetica", bg="transparent", pointsize=10)
26    qcc(mean, type="xbar.one", std.dev="MR", labels=studies, title=title,
27      ylab=efect, xlab="",
28      axes.las=2, nsigmas=3)
29    dev.off()
30    return(TRUE)
31  }

32  #-----
33  #-- Operacionalizaciones de Causa
34  C_Op_A <- "T. Insp. A"
35  C_Op_B <- "T. Insp. B"

36  #-- Operacionalizaciones de Efecto
37  E_Op <- "Percentage of Observed faults"

38

39  ds <- read.table("part01/res_comp/exp_res_part01.txt", header=T)

```



```
40 ds$J <- 1-(3/(4*(ds$n.e+ds$n.c-2)-1))
41 ds$Swithin <- sqrt( ((ds$n.e-1) * ds$sd.e^2 + (ds$n.c-1) * ds$sd.c^2) / (
    ds$n.e+ds$n.c-2) )
42 ds$d <- (ds$mean.e - ds$mean.c) / ds$Swithin
43 ds$g <- ds$J * ds$d
44 ds$Vd <- ( (ds$n.e + ds$n.c)/(ds$n.e * ds$n.c) ) + ds$d^2/(2*(ds$n.e +
    ds$n.c))
45 ds$Vg <- ds$J^2*ds$Vd
46 ds$SEg <- sqrt(ds$Vg)

48 meta01 <- metagen(ds$g, ds$SEg, sm="SMD", data=ds, studlab=paste(studlab),
    label.e=C_0p_A,
49     label.c=C_0p_B, complab=paste(C_0p_A," vs. ",C_0p_B), outclab=E_0p )

51 meta01$n.e<-addvar(meta01,ds,"n.e")
52 meta01$mean.e<-addvar(meta01,ds,"mean.e")
53 meta01$sd.e<-addvar(meta01,ds,"sd.e")
54 meta01$n.c<-addvar(meta01,ds,"n.c")
55 meta01$mean.c<-addvar(meta01,ds,"mean.c")
56 meta01$sd.c<-addvar(meta01,ds,"sd.c")

58 #-- Genera diagrama de bosque
59 PrintMeta("part01/res_comp/_meta_part01.eps", meta01)

61 #-- Genera gráficos de control
62 PrintCC("part01/res_comp/_cc_a_part01.eps", ds$mean.e, ds$studlab, paste("
    Control Chart for ",C_0p_A), E_0p)
63 PrintCC("part01/res_comp/_cc_b_part01.eps", ds$mean.c, ds$studlab, paste("
    Control Chart for ",C_0p_B), E_0p)
```

Listado C.3: r_script02

Apéndice D

Formularios del Procedimiento para Comparar Replicaciones

RESUMEN: En este apéndice se encuentran disponibles los formularios usados en el procedimiento para comparar replications.

En este apéndice se encuentran disponibles los formularios usados en el procedimiento para comparar replications. A continuación se listan los formularios utilizados.

- Formulario de Constructos de Causa y Efecto (PCR-1)
- Formulario de Operacionalizaciones de Causa y Efecto (PCR-2)
- Formulario de Agrupaciones de Constructos y Operacionalizaciones (PCR-3)
- Formulario de Registro de la Situación Experimental por Experimento (PCR-4)
- Formulario de Registro de la Tabla Maestra de Cambios (PCR-5)
- Formulario de Registro para Tablas de Particiones (PCR-6)
- Formulario de Registro para la Tabla de Resultados Experimentales (PCR-7)
- Formulario de Registro de Posibles Causas que Influyen en los Resultados de la Partición (PCR-8)
- Formulario de Registro del Análisis de Resultados Experimentales (PCR-9)
- Formulario de Registro de Variaciones en Situación Experimental (PCR-10)
- Formulario de Registro de Pieza de Conocimiento para Profesional (PCR-11)
- Formulario de Registro de Pieza de Conocimiento para Investigador (PCR-12)

D.1 Formulario de Constructos de Causa y Efecto (PCR-1)

Investigador:

Fecha:

Objeto de Estudio:

Constructos		
ID Experimento	Constructos de Causa	Constructos de Efecto

D.2 Formulario de Operacionalizaciones de Causa y Efecto (PCR-2)

Investigador:

Fecha:

Objeto de Estudio:

	Cons. de Causa		Cons. de Efecto	
ID Exp.	Op. Causa _i	Op. Causa _{i+1}	Op. Efecto _j	Op. Efecto _{j+1}

D.3 Formulario de Agrupaciones de Constructos y Operacionalizaciones (PCR-3)

Investigador:

Fecha:

Objeto de Estudio:

Agrupaciones de Constructos y Operacionalizaciones				
ID Exp.	Con. Efecto	Op. Efecto	Par Con. Causa	Par Op. Causa

D.4 Formulario de Registro de la Situación Experimental por Experimento (PCR-4)

Investigador:

Fecha:

Experimento:

Experimentadores y Sitio		
	Elemento Principal	Descripción
Experimentadores	Diseñador	
	Entrenador	
	Monitor	
	Medición	
	Analista	
Sitio	Sitio	

Protocolo Experimental		
Elemento Principal	Sub-elemento	Descripción
Análisis	Técnica de análisis de datos	
Diseño	Diseño experimental	
	Orden de aplicación de los tratamientos	
	Combinación del entrenamiento con las sesiones	
	Tratamientos por sesión	
	Sesiones requeridas por sujeto para aplicar el tratamiento	
	Tareas por sujeto por sesión	
Instrumentos	Instrumentos o equipos de medición	
Objetos experimentales	Propiedad _i	
	Propiedad _{i+1}	
Materiales	Formularios del experimento	
	Instrucciones del experimento	
	Guías suplementarias	

Operacionalizaciones de los Constructos		
Elemento Principal	Sub-elemento	Descripción
(Causas) Tratamiento _i	Tratamiento acorde a la literatura	
	Procedimiento de aplicación del tratamiento	
	Tratamiento como instrucciones	
	Recursos del tratamiento	
	Aspectos de transmisión del tratamiento	
Tratamiento _{i+1}	Tratamiento acorde a la literatura	
	Procedimiento de aplicación del tratamiento	
	Tratamiento como instrucciones	
	Recursos del tratamiento	
	Aspectos de transmisión del tratamiento	
(Efecto) Constructo _i	Métrica de la medida _i	
	Procedimiento de medición	
Constructo _{i+1}	Métrica de la medida _{i+1}	
	Procedimiento de medición	

Propiedades de las Poblaciones		
Elemento Principal	Sub-elemento	Descripción
Propiedades de los sujetos	Propiedad _{<i>i</i>}	
	Propiedad _{<i>i</i>+1}	
Propiedades los objetos	Propiedad _{<i>j</i>}	
	Propiedad _{<i>j</i>+1}	

D.5 Formulario de Registro de la Tabla Maestra de Cambios (PCR-5)

Investigador:

Fecha:

Experimentos de la tabla maestra:

Experimentadores y Sitio			
	Elemento Principal	Experimento _i	Experimento _{i+1}
Experimentadores	Diseñador		
	Entrenador		
	Monitor		
	Medición		
	Analista		
Sitio	Sitio		

Protocolo Experimental			
Elemento Principal	Sub-elemento	Exp _i	Exp _{i+1}
Análisis	Técnica de análisis de datos		
Diseño	Diseño experimental		
	Orden de aplicación de los tratamientos		
	Combinación del entrenamiento con las sesiones		
	Tratamientos por sesión		
	Sesiones requeridas por sujeto para aplicar el tratamiento		
	Tareas por sujeto por sesión		
Instrumentos	Instrumentos o equipos de medición		
Objetos experimentales	Propiedad _i		
	Propiedad _{i+1}		
Materiales	Formularios del experimento		
	Instrucciones del experimento		
	Guías suplementarias		

Operacionalizaciones de los Constructos			
Elemento Principal	Sub-elemento	Exp _i	Exp _{i+1}
(Causas) Tratamiento _i	Tratamiento acorde a la literatura		
	Procedimiento de aplicación del tratamiento		
	Tratamiento como instrucciones		
	Recursos del tratamiento		
	Aspectos de transmisión del tratamiento		
Tratamiento _{i+1}	Tratamiento acorde a la literatura		
	Procedimiento de aplicación del tratamiento		
	Tratamiento como instrucciones		
	Recursos del tratamiento		
	Aspectos de transmisión del tratamiento		
(Efecto) Constructo _i	Métrica de la medida _i		
	Procedimiento de medición		
Constructo _{i+1}	Métrica de la medida _{i+1}		
	Procedimiento de medición		

Propiedades de las Poblaciones			
Elemento Principal	Sub-elemento	Exp_i	Exp_{i+1}
Propiedades de los sujetos	Propiedad _{i}		
	Propiedad _{$i+1$}		
Propiedades los objetos	Propiedad _{j}		
	Propiedad _{$j+1$}		

D.6 Formulario de Registro para Tablas de Particiones (PCR-6)

Investigador:

Fecha:

Id Partición:

Partición:

Experimentadores y Sitio					
	Elemento Principal	ID	Categorías	Exp_i	Exp_{i+1}
Experimentadores	Diseñador	1.1			
	Entrenador	1.2			
	Monitor	1.3			
	Medición	1.4			
	Analista	1.5			
Sitio	Sitio	2			

Protocolo Experimental					
Elemento Principal	Sub-elemento	ID	Categorías	Exp _i	Exp _{i+1}
Análisis	Técnica de análisis de datos	3.1.1			
Diseño	Diseño experimental	3.2.1			
	Orden de aplicación de los tratamientos	3.2.2			
	Combinación del entrenamiento con las sesiones	3.2.3			
	Tratamientos por sesión	3.2.4			
	Sesiones requeridas por sujeto para aplicar el tratamiento	3.2.5			
	Tareas por sujeto por sesión	3.2.6			
Instrumentos	Instrumentos o equipos de medición	3.3.1			
Objetos experimentales	Propiedad _i	3.4.1			
	Propiedad _{i+1}	3.4.2			
Materiales	Formularios del experimento	3.5.1			
	Instrucciones del experimento	3.5.2			
	Guías suplementarias	3.5.3			

Operacionalizaciones de los Constructos					
Elemento Principal	Sub-elemento	ID	Categorías	Exp _i	Exp _{i+1}
(Causas) Tratamiento _i	Tratamiento acorde a la literatura	4.1.1.1			
	Procedimiento de aplicación del tratamiento	4.1.1.2			
	Tratamiento como instrucciones	4.1.1.3			
	Recursos del tratamiento	4.1.1.4			
	Aspectos de transmisión del tratamiento	4.1.1.5			
Tratamiento _{i+1}	Tratamiento acorde a la literatura	4.1.2.1			
	Procedimiento de aplicación del tratamiento	4.1.2.2			
	Tratamiento como instrucciones	4.1.2.3			
	Recursos del tratamiento	4.1.2.4			
	Aspectos de transmisión del tratamiento	4.1.2.5			
(Efecto) Constructo _i	Métrica de la medida _i	4.2.1			
	Procedimiento de medición	4.2.2			

Propiedades de las Poblaciones					
Elemento Principal	Sub-elemento	ID	Categorías	Exp_i	Exp_{i+1}
Propiedades de los sujetos	Propiedad _{i}	5.1.1			
	Propiedad _{$i+1$}	5.1.2			
Propiedades los objetos	Propiedad _{j}	5.2.1			
	Propiedad _{$j+1$}	5.2.2			

D.7 Formulario de Registro para la Tabla de Resultados Experimentales (PCR-7)

Investigador:

Fecha:

Id Partición:

Partición:

	Tratamiento A			Tratamiento B		
ID Exp.	n_1	\bar{x}_1	s_1	n_2	\bar{x}_2	s_2

D.8 Formulario de Registro de Posibles Causas que Influyen en los Resultados de la Partición (PCR-8)

Investigador:

Fecha:

Id Partición:

Partición:

ID Experimento	Cambio que se cree influye en resultados

D.9 Formulario de Registro del Análisis de Resultados Experimentales (PCR-9)

Investigador:

Fecha:

Información de la Partición	
ID Partición	
Partición	
Replicaciones de la partición	
Información del tipo de resultados obtenidos	
Resultados del análisis de similitudes	
Tipo de resultados encontrados	
Información de los resultados	
Tipo de inconsistencia observada	
Información de resultados consistentes	
Replicaciones consideradas para la pieza de conocimiento	
Comparación del par de tratamientos	
Tipo de efecto global	
Tamaño del efecto global	
Análisis de Causas	
Nuevas hipótesis de trabajo	

D.10 Formulario de Registro de Variaciones en Situación Experimental (PCR-10)

Investigador:

Fecha:

Información de la partición					
ID Partición					
Partición					
Replicaciones incluidas			Replicaciones apartadas		
Variaciones en la situación experimental					
Sección	Elemento	(sub) elementos variados	total de elementos	% de variación	% de var. por secc.
Experimentadores	Diseñador				
	Entrenador				
	Monitor				
	Medición				
	Analista				
Sitio					
Protocolo experimental	Análisis				
	Diseño				
	Instrumentos				
	Objetos experimentales				
	Materiales				
Operacionalizaciones de los constructos	Tratamiento A				
	Tratamiento B				
	Efecto				
Poblaciones	Sujetos				
	Objetos experimentales				

D.11 Formulario de Registro de Pieza de Conocimiento para Profesional (PCR-11)

Investigador:

Fecha:

Información de la partición		
ID Partición		
Partición		
Replicaciones de la partición		
Resultados del análisis de similitudes		
Tipo de resultados encontrados		
Tipo de inconsistencia observada		
Información de la comparación		
Replicaciones consideradas para la pieza de conocimiento		
Comparación		
Tipo de efecto global		
Tamaño del efecto global		
Variaciones en la situación experimental		
Secciones	Porcentaje de variaciones	Tipo de evento observado
Experimentadores		
Sitio		
Protocolo experimental		
Operacionalizaciones de los constructos		
Poblaciones		
Poder de confirmación de resultados		
Confiabilidad de la pieza de conocimiento		

D.12 Formulario de Registro de Pieza de Conocimiento para Investigador (PCR-12)

Investigador:

Fecha:

Información de la partición			
ID Partición			
Partición			
Replicaciones de la partición			
Resultados del análisis de similitudes			
Tipo de resultados encontrados			
Información de los resultados			
Tipo de inconsistencia observada			
Información de la comparación			
Replicaciones consideradas para la pieza de conocimiento			
Comparación			
Tipo de efecto global			
Tamaño del efecto global			
Variaciones en la situación experimental			
Secciones	Porcentaje de variaciones	Tipo de evento observado	Tipos de Replicaciones
Experimentadores			
Sitio			
Protocolo experimental			
Operacionalizaciones de los constructos			
Poblaciones			
Poder de confirmación de resultados			
Confiabilidad de la pieza de conocimiento			
Análisis de Causas			
Nuevas hipótesis de trabajo			

Apéndice E

Descripción General de las Técnicas de Evaluación de Software

En este apéndice se describe de manera general las técnicas de evaluación de software empleadas en el conjunto de replicaciones usado en la evaluación de esta investigación. Para más información de estas técnicas consultar los trabajos de Howden (Howden, 1980), Myers (Myers, 1979), Marick (Marick, 1995) y Linger et al. (Linger et al., 1979).

E.1 Lectura de Código por Abstracciones Sucesivas

La lectura por abstracciones sucesiva sirve para inspeccionar código, y no otro tipo de artefacto como requisitos o diseño. La idea sobre la que se basa esta técnica de lectura para detectar defectos es en la comparación entre la especificación del programa (es decir, el texto que describe lo que el programa debería hacer) y lo que el programa hace realmente. Naturalmente, todos aquellos puntos donde no coincida lo que el programa debería hacer con lo que el programa hace es un defecto.

Dado que comparar código con texto (la especificación) es inapropiado pues se estarían comparando unidades distintas, se hace necesario convertir ambos artefactos a las mismas “unidades”. Lo que se hace, entonces, es convertir el programa en una especificación en forma de texto. De modo que podamos comparar especificación (texto) con especificación (texto).

Obtener una especificación a partir de un código significa recorrer el camino de la programación en sentido inverso. El sentido directo (cuando se programa) es obtener un código a partir de una especificación. Este camino se recorre en una serie de pasos (que a menudo quedan ocultos en la mente del programador y no se hacen explícitos, pero que no por eso dejan de existir). El recorrido directo del camino de la especificación al código consta de los siguientes pasos:

1. Leer la especificación varias veces hasta que el programador ha entendido lo que el código debe hacer.
2. Descomponer la tarea que el programa debe hacer en subtarefas, que típicamente se corresponderán con las funciones o módulos que compondrán el programa. Esta descomposición muestra la relación entre funciones, que no siempre es secuencial, sino típicamente en forma de árbol: funciones alternativas que se ejecutarán dependiendo de alguna condición; Funciones suplementarias que se ejecutarán siempre si una ejecuta la otra; etc.

3. Para cada uno de estas subtareas (funciones):

- (a) Hacer una descripción sistemática (típicamente en pseudocódigo) de cómo realizar la tarea. En esta descripción se pueden ya apreciar las principales estructuras de las que constará la función (bucles, condicionales, etc.)
- (b) Programar cada línea de código que compone la función

Como puede observarse en este proceso, el programador trabaja desde la especificación por descomposiciones sucesivas. Es decir, dividiendo una idea compleja (la especificación) en cosas cada vez más sencillas (primero las funciones, luego las estructuras elementales de los programas, finalmente las líneas de código). Este tipo de tarea se realiza de arriba hacia abajo, partiendo de la especificación (arriba o nodo raíz) y llegando a n líneas de código (abajo o nodos hojas).

Si se desea obtener una especificación a partir de un código se debe hacer este mismo recorrido pero en sentido contrario: de abajo hacia arriba. Por tanto, se debe comenzar con las líneas de código, agruparlas en estructuras elementales, y éstas en funciones y éstas en un todo que será la descripción del comportamiento del programa. En este caso no se está descomponiendo, sino componiendo. La tarea que se realiza es de abstracción. De ahí el nombre de la técnica: abstracción sucesiva (ir sucesivamente –ascendiendo en niveles cada vez superiores– abstrayendo qué hace cada elemento primero las líneas, luego las estructuras, finalmente las funciones).

Esta técnica requiere que el inspector lea una serie de líneas de código y que abstraiga la función que estas líneas computan. El inspector debe repetir este procedimiento hasta que la función final del código que se está inspeccionando se haya abstraído y pueda compararse con la especificación original del programa.

Más concretamente, el proceso que se debe seguir para realizar la abstracción sucesiva es el siguiente:

1. Leer de forma general el código para tener una idea general del programa. Determinar las dependencias entre las funciones individuales del código fuente (quién llama a quién). Para esto puede usarse un árbol de llamadas para representar tales dependencias comenzando por las hojas (funciones que no llaman a nadie) y acabando por la raíz. (función principal).
2. Comprender qué hace cada función. Para ello se deberá: Entender la estructura de cada función individual identificando las estructuras elementales (secuencias, condicionales, bucles, etc.) y marcándolas; Combinar las estructuras elementales para formar estructuras más grandes hasta que se haya entendido la función entera. Es decir, para cada función y comenzando desde las funciones hoja y acabando por la raíz:
 - (a) Identificar las estructuras elementales de cada función y marcarlas de la más interna a la más externa.
 - (b) Determinar el significado de cada estructura comenzando con la más interna. Para ello pueden seguirse las siguientes recomendaciones: Usar los números de línea para identificar las líneas que abarca una estructura e indicar a su lado qué hace. Evitar utilizar conocimiento implícito que no resida en la estructura (valores iniciales, entradas o valores de parámetros). Usar principios generalmente aceptados del dominio de aplicación para mantener la descripción breve y entendible (“búsqueda en profundidad” en lugar de describir lo que hace la búsqueda en profundidad).

- (c) Especificar el comportamiento de la función entera. Es decir, utilizar la información de los pasos 2a y 2b para entender qué hace la función y describirlo en texto.
- 3. Combinar el comportamiento de cada función y las relaciones entre ellas para entender el comportamiento del programa entero. El comportamiento deberá describirse en texto. Esta descripción es la especificación del programa obtenida a partir del código.
- 4. Comparar la especificación obtenida con la especificación original del programa. Cada punto donde la especificación original y especificación generada a partir del código no coincidan es un defecto. Anotar los defectos en una lista.

E.2 Pruebas de Caja Blanca o Estructurales

A este tipo de técnicas se le conoce también como técnicas de caja transparente o de cristal. Este método se centra en cómo diseñar los casos de prueba atendiendo al comportamiento interno y la estructura del programa. Se examina así la lógica interna del programa sin considerar los aspectos de rendimiento.

El objetivo de la técnica es diseñar casos de prueba para que se ejecuten, al menos una vez, todas las sentencias del programa, y todas las condiciones tanto en su vertiente verdadera como falsa. Como se ha indicado ya, puede ser impracticable realizar una prueba exhaustiva de todos los caminos de un programa. Por ello se han definido distintos criterios de cobertura lógica, que permiten decidir qué sentencias o caminos se deben examinar con los casos de prueba. Estos criterios son:

- *Cobertura de sentencias.* Se escriben casos de prueba suficientes para que cada sentencia en el programa se ejecute, al menos, una vez.
- *Cobertura de decisión.* Se escriben casos de prueba suficientes para que cada decisión en el programa se ejecute una vez con resultado verdadero y otra con el falso.
- *Cobertura de condiciones.* Se escriben casos de prueba suficientes para que cada condición en una decisión tenga una vez resultado verdadero y otra falso.
- *Cobertura decisión/condición.* Se escriben casos de prueba suficientes para que cada condición en una decisión tome todas las posibles salidas, al menos una vez, y cada decisión tome todas las posibles salidas, al menos una vez.
- *Cobertura de condición múltiple.* Se escriben casos de prueba suficientes para que todas las combinaciones posibles de resultados de cada condición se invoquen al menos una vez.
- *Cobertura de caminos.* Se escriben casos de prueba suficientes para que se ejecuten todos los caminos de un programa. Entendiendo camino como una secuencia de sentencias encadenadas desde la entrada del programa hasta su salida.

E.3 Pruebas de Caja Negra o Funcionales

También conocidas como Pruebas de Comportamiento, estas pruebas se basan en la especificación del programa o componente a ser probado para elaborar los casos de prueba. El componente se ve como una “Caja Negra” cuyo comportamiento sólo puede ser determinado estudiando sus entradas y las salidas obtenidas a partir de ellas. No obstante, como

el estudio de todas las posibles entradas y salidas de un programa sería impracticable se selecciona un conjunto de ellas sobre las que se realizan las pruebas. Para seleccionar el conjunto de entradas y salidas sobre las que trabajar, hay que tener en cuenta que en todo programa existe un conjunto de entradas que causan un comportamiento erróneo en nuestro sistema, y como consecuencia producen una serie de salidas que revelan la presencia de defectos. Entonces, dado que la prueba exhaustiva es imposible, el objetivo final es pues, encontrar una serie de datos de entrada cuya probabilidad de pertenecer al conjunto de entradas que causan dicho comportamiento erróneo sea lo más alto posible.

Al igual que ocurría con las técnicas de Caja Blanca, para confeccionar los casos de prueba de Caja Negra existen distintos criterios. Algunos de ellos son:

- *Particiones de Equivalencia.* Una clase de equivalencia representa un conjunto de estados válidos o no válidos para condiciones de entrada. Típicamente, una condición de entrada es un valor numérico específico, un rango de valores, un conjunto de valores relacionados o una condición lógica.
- *Análisis de Valores Límite.* los errores tienden a darse más en los límites del campo de entrada que en el centro. El análisis de valores límite lleva a una elección de casos de prueba que ejerciten los valores límite.

El análisis de valores límite es una técnica de diseño de casos de prueba que completa a la partición equivalente. En lugar de seleccionar cualquier elemento de una clase de equivalencia, el BVA (*Boundary Value Analysis*, Análisis de Valores Límite) lleva a la elección de casos de prueba en los extremos de la clase. En lugar de centrarse solamente en las condiciones de entrada, el BVA obtiene casos de prueba también para el campo de salida.

- *Métodos Basados en Grafos.* en este método se debe entender los objetos (objetos de datos, objetos de programa tales como módulos o colecciones de sentencias del lenguaje de programación) que se modelan en el software y las relaciones que conectan a estos objetos. Una vez que se ha llevado a cabo esto, el siguiente paso es definir una serie de pruebas que verifiquen que todos los objetos tienen entre ellos las relaciones esperadas. En este método, 1) se crea un grafo de objetos importantes y sus relaciones, y 2) se diseña una serie de pruebas que cubran el grafo de manera que se ejerciten todos los objetos y sus relaciones para descubrir errores.
- *Gráfico de Causa-Efecto.* La gráfica Causa-efecto representa una ayuda gráfica en seleccionar, de una manera sistemática, un gran conjunto de casos de prueba. Tiene un efecto secundario beneficioso en precisar estados incompletos y ambigüedades en la especificación.

Un gráfico de causa-efecto es un lenguaje formal al cual se traduce una especificación. El gráfico es realmente un circuito de lógica digital (una red combinatoria de lógica), pero en vez de la notación estándar de la electrónica, se utiliza una notación algo más simple. No hay necesidad de tener conocimiento de electrónica con excepción de una comprensión de la lógica booleana (entendiendo los operadores de la lógica y, o, y no).

Apéndice F

Materiales Empleados en Replicaciones UPM06..10

En este apéndice se encuentra disponible el material usado en las replicaciones *upm06* a *upm10*. En estas replicaciones se aplicaron tres técnicas de evaluación de software (lectura de código por abstracciones sucesivas, prueba funcional o caja negra y prueba estructural o de caja blanca). Las técnicas se usan para examinar tres programas software. Este material se compone de instrucciones, formularios, guías suplementarias, especificaciones y los códigos fuente de los programas usados en estas replicaciones. A continuación se listan los diferentes elementos de este material.

Para la técnica de revisión de código por abstracciones sucesivas:

- Instrucciones de aplicación de la técnica (ver F.1)
- Formulario de recogida de datos (ver F.2)
- Formulario de abstracciones (ver F.3)
- Formulario de inconsistencias (ver F.4)
- Formulario de conclusiones (ver F.5)

Para la técnica de prueba funcional:

- Instrucciones de aplicación de la técnica (ver F.6)
- Formulario de recogida de datos (ver F.7)
- Formulario de clases de equivalencia (ver F.8)
- Formulario de datos de prueba (ver F.9)
- Formulario de salidas (ver F.10)
- Formulario de fallos (ver F.11)
- Formulario de conclusiones (ver F.12)

Para la técnica de prueba estructural:

- Instrucciones de aplicación de la técnica (ver F.13)
- Formulario de recogida de datos (ver F.14)
- Formulario de datos de prueba (ver F.15)
- Formulario de salidas (ver F.16)
- Formulario de fallos (ver F.17)
- Formulario de conclusiones (ver F.18)

Programa *cmdline*:

- Código fuente (ver F.19)
- Especificación (ver F.20)
- Suplemento técnica de revisión de código (ver F.21)
- Suplemento técnica de prueba funcional (ver F.22)
- Suplemento técnica de prueba estructural (ver F.23)

Programa *nametbl*:

- Código fuente (ver F.24)
- Especificación (ver F.25)
- Suplemento técnica de revisión de código (ver F.26)
- Suplemento técnica de prueba funcional (ver F.27)
- Suplemento técnica de prueba estructural (ver F.28)

Programa *ntree*:

- Código fuente (ver F.29)
- Especificación (ver F.30)
- Suplemento técnica de revisión de código (ver F.31)
- Suplemento técnica de prueba funcional (ver F.32)
- Suplemento técnica de prueba estructural (ver F.33)

F.1 Instrucciones para aplicar la técnica “Revisión de Código” (E10)

Consultar la hoja suplementaria para aplicar la técnica “Revisión de Código” con el Programa ‘X’ a medida que se leen las instrucciones

Lectura de código

1. Escribe el nombre en el Formulario de abstracciones (E12) y en el de recogida de datos (E11).
2. Lee por encima el código para tener una idea general del componente. Si te parece descubrir faltas mientras haces este paso a alguno de los dos siguientes, márcalas. No obstante, no pierdas demasiado tiempo en hacer un análisis preciso de ellas.
3. Determina las dependencias entre las funciones individuales del código fuente, usando para ello un árbol de llamadas. Normalmente las funciones ya estarán ordenadas en el código, de tal modo que las funciones de bajo nivel (hojas) están al comienzo y las funciones de más alto nivel (raíz) aparecen al final. Comienza aplicando la técnica de revisión de código con las funciones hoja y sigue hacia la raíz.
4. Intenta entender la estructura de cada función individual identificando las estructuras elementales (secuencias condicionales, bucles) y marcándolas. Combina las estructuras elementales para formar estructuras más grandes hasta que hayas entendido la función entera.
5. Trata de determinar el significado de cada estructura comenzando con la más interna y anota el resultado en el formulario para tal propósito (E12). Usa números de línea (línea x-y) cuando hagas eso. Evita utilizar conocimiento implícito que no resida en la estructura, por ejemplo valores iniciales, entradas o valores de parámetros. Describe las partes lo más funcionalmente posible. Mientras hagas eso, usa principios generalmente aceptados del dominio de aplicación para mantener la descripción breve y entendible. Por ejemplo, en el caso de búsqueda en un árbol, menciona “búsqueda en profundidad” en lugar de describir lo que hace la búsqueda en profundidad.
6. Cuando hayas completado las abstracciones, incluyendo tu versión de la especificación del programa, deberás “congelarlas”. Pasa al siguiente paso.

Búsqueda de inconsistencias

7. Recoge el Formulario para inconsistencias (E13) y la especificación del código (E01). Pon el nombre al Formulario para inconsistencias.
8. Lee con detenimiento la especificación que se te ha entregado. Dicha especificación describe el componente como un todo, no las funciones individuales. Detecta las posibles inconsistencias comparando la especificación con tu descripción del componente. Escribe las inconsistencias que hayas detectado en el formulario E13.

Para ello, numera las inconsistencias que hayas encontrado desde 1 hasta n en la columna denominada “Nº Inc” (número de inconsistencia) del formulario para inconsistencias. Describe la inconsistencia en las columnas a la derecha del número de inconsistencia.
9. Una vez creas que has detectado todas las inconsistencias, entrega todo el material a la persona a cargo del ejercicio. Ya has terminado.

F.2 Formulario de Recogida de datos para la Revisión de Código (E11)

Nombre: _____

Aspectos de tiempo

Para cada entrada referida al tiempo que ha transcurrido, deduce el tiempo que hayas tomado para descansos, etc.

1. ¿A qué hora comenzaste el ejercicio de “revisión de código”? (hora: minutos)
2. ¿Cuánto tiempo necesitaste para construir las abstracciones? (horas:minutos)
3. ¿Cuántos niveles de abstracciones has generado?
4. ¿A qué hora comenzaste a buscar las inconsistencias? (hora:minutos)
5. ¿Cuánto tiempo necesitaste para encontrar las inconsistencias? (horas:minutos)
6. ¿A qué hora terminaste el ejercicio? (hora:minutos)
7. ¿Podrías asegurar que encontraste todas las faltas? Estima el porcentaje de faltas que has encontrado (en
8. ¿Cómo de bien crees que has efectuado la revisión de código? Marca la escala apropiadamente.

Estimación	0	1	2	3	4	5
Comparación	fatal	bastan- te mal	regular	bien	muy bien	perfecta- mente

**F.3 Formulario de abstracciones para la Revisión de Código
 (E12)**

Identificador (Nombre y Grupo): _____

Abstracciones	
Líneas	Abstracción

Inconsistencias detectadas				
Nº Inc.	Comportamiento esperado	Número de líneas (comienzo, fin) de la abstracción y explicación breve de la inconsistencia	Seguridad de que es una inconsistencia (%)	Encontrada con la técnica (si/no)

F.5 Formulario de Conclusiones para la Revisión de Código (E14)

Nombre: _____

Sobre la preparación del ejercicio

1. ¿Cuánto tiempo has dedicado a estudiar la técnica de revisión de código? (Indica el número de horas)
2. ¿Cuántos ejercicios has hecho por tu cuenta (fuera de clase) para la técnica de revisión de código? Cuantifícalo con un número incluyendo los ejercicios obligatorios si los hubo.

Sobre la aplicación de la técnica

3. Indica y numera los pasos que has seguido para aplicar la técnica de revisión de código.
4. ¿Cuánto de seguro estás de que has detectado todas las faltas?

Muy seguro	Medianamente seguro	Poco seguro	Nada seguro
------------	---------------------	-------------	-------------
5. ¿Has encontrado alguna situación en la que no hayas sido capaz de determinar si estabas frente a una falta o no? Si En caso afirmativo, explica la situación.

Sobre la visibilidad de las faltas

6. ¿Ha sido más fácil detectar las faltas en algún trozo de código en particular?

F.6 Instrucciones para aplicar la técnica de “Pruebas Funcionales” (E30)

Consultar la hoja suplementaria para aplicar la técnica “Pruebas Funcionales” al Programa X a medida que se leen las instrucciones

Generación de casos de prueba

1. Pon el nombre al formulario para las clases de equivalencia (E32), al de datos de prueba para las pruebas funcionales (E33), y al de recogida de datos (E31).
2. Lee cuidadosamente la especificación del código. Úsala para obtener las clases de equivalencia y escríbelas en el formulario E32.
3. Genera casos de prueba eligiendo valores de las clases de equivalencia y anota los casos de prueba en el formulario E33. La columna nombrada “Nº Clase Equival.” Debería contener el/los número/s de las clases de equivalencia del formulario E32 que el caso de prueba está ejercitando. Mientras hagas esto, ignora cualquier dependencia del sistema (no del programa).
4. En este momento se congelarán los casos de prueba generados. No puedes generar ningún caso de prueba más.

Ejecución de los casos de prueba

5. Localiza el fichero ejecutable correspondiente al programa, pide el formulario para anotar las salidas del programa (E34) y pon el nombre.
6. Lee los casos de prueba y determina (según las especificaciones) cuál debería ser la salida esperada. A continuación, ejecuta los casos de prueba que has generado y toma nota de la salida observada en el formulario E34.

Identificación de fallos

7. Recoge el formulario de fallos para las Pruebas Funcionales (E35) y no olvides poner tu nombre.
8. Examina los valores esperados y los obtenidos con detenimiento. Trata de encontrar posibles fallos comparando los resultados esperados (de acuerdo a la especificación) con la salida de tus casos de prueba. Anota cualquier discrepancia entre las dos salidas en el formulario E35. Por favor, numera los fallos que hayas encontrado de 1 a n en la columna denominada “Nº de fallo”.

Conclusiones

9. Recoge el formulario de conclusiones para las pruebas funcionales (E36).
10. Cuando creas que has detectado todos los fallos, entrega todos los formularios a la persona a cargo del ejercicio. Ya has terminado.

F.7 Formulario de Recogida de datos para las Pruebas Funcionales (E31)

Nombre: _____

Aspectos de tiempo

Para cada entrada referida al tiempo que ha transcurrido, deduce el tiempo que hayas tomado para descansos, etc.

1. ¿A qué hora comenzaste el ejercicio de “pruebas funcionales”?
(hora:minutos)
2. ¿Cuántas clases de equivalencia generaste?
3. ¿Cuántos casos de prueba generaste?
4. ¿Cuánto tiempo necesitaste para generar los casos de prueba? (horas:minutos)
5. ¿Cuánto tiempo necesitaste para teclear y ejecutar los casos de prueba? (hora:minutos)
6. ¿A qué hora comenzaste a buscar los fallos? (hora:minutos)
7. ¿Cuánto tiempo necesitaste para encontrar los fallos? (horas:minutos)
8. ¿A qué hora terminaste el ejercicio? (hora:minutos)
9. ¿Podrías asegurar que encontraste todas las fallos? Estima el porcentaje de fallos que has encontrado (en
10. ¿Cómo de bien crees que has efectuado las pruebas funcionales? Marca la escala apropiadamente.

Estimación	0	1	2	3	4	5
Comparación	fatal	bastan- te mal	regular	bien	muy bien	perfecta- mente

F.8 Formulario de Clases de Equivalencia para las Pruebas Funcionales (E32)

Nombre: _____

Clases de Equivalencia	
Número	Descripción

F.9 Formulario de Datos de Prueba para las Pruebas Funcionales (E33)

Nombre: _____

Nº de caso	Nº Clase Equivalencia	Datos de prueba

F.10 Formulario de Salidas para las Pruebas Funcionales (E34)

Nombre: _____

Nº de caso	Salida Esperada	Salida Observada

F.11 Formulario de fallos para las Pruebas Funcionales (E35)

Nombre: _____

Fallos detectados				
Nº de fallo.	Nº de caso de prueba	Breve descripción del fallo	Seguridad de que es un fallo (%)	Encontrado con la técnica (si/no)

F.12 Formulario de Conclusiones para las Pruebas Funcionales (E36)

Nombre: _____

Sobre la preparación del ejercicio

1. ¿Cuánto tiempo has dedicado a estudiar la técnica funcional? (Indica el número de horas)
2. ¿Cuántos ejercicios has hecho por tu cuenta (fuera de clase) para la técnica funcional? Cuantifícalo con un número incluyendo los ejercicios obligatorios si los hubo.

Sobre la generación de los casos de prueba

3. Indica y numera los pasos que has seguido para aplicar la técnica funcional.
4. ¿Cuánto de seguro estás de que los casos de prueba que has generado ejercitan todas las faltas?

Muy seguro	Medianamente seguro	Poco seguro	Nada seguro
------------	---------------------	-------------	-------------

Sobre la ejecución de los casos de prueba

5. ¿Cuánto de seguro estás de que has detectado todos los fallos que se han revelado durante la ejecución de los casos de prueba?

Muy seguro	Medianamente seguro	Poco seguro	Nada seguro
------------	---------------------	-------------	-------------

6. ¿Has encontrado alguna situación en la que no hayas sido capaz de determinar si estabas frente a un fallo o no? Si En caso afirmativo, explica la situación.

F.13 Instrucciones para aplicar la técnica “Pruebas Estructurales” (E20)

Consultar la hoja suplementaria para aplicar la técnica “Pruebas Estructurales” al Programa X a medida que se leen las instrucciones

Generación de casos de prueba

1. Escribe el nombre en el formulario de datos de prueba (E22) y en el de recogida de datos (E21)
2. Lee de pasada el código para tener una idea general del componente. Si te parece descubrir faltas mientras haces este paso a alguno de los dos siguientes, márcalas. No obstante, no pierdas demasiado tiempo en hacer un análisis preciso de ellas.
3. Comienza a generar los casos de prueba para cumplir el criterio de cobertura (que aparece explicado al final de estas instrucciones). En la hoja suplementaria aparecen propiedades especiales del componente.
4. Anota el propósito del caso de prueba (una breve descripción de la prueba) y el caso de prueba en el formulario E22.
5. Una vez que hayas alcanzado la cobertura deseada o de que no puedes conseguir una mejor, has terminado la primera parte de este ejercicio. Por favor, no generes más casos de prueba a partir de éste momento.

Ejecución de casos de prueba

6. Localiza el fichero ejecutable del programa correspondiente, recoge el formulario para anotar las salidas del programa (E23) y pon el nombre.
7. Ejecuta los casos de prueba que has generado y anota la salida obtenida en el formulario E23.

Identificación de fallos

8. Recoge el formulario de fallos para las pruebas estructurales (E24) y la especificación del código (E01). Escribe el nombre al formulario de fallos (E24).
9. Lee la especificación del componente y busca los posibles fallos comparando los resultados esperados (de acuerdo con la especificación) con las salidas que se observaron y que están anotadas en el formulario E23.
10. Anota los fallos encontrados en el formulario E24. Por favor, numera las fallos que has encontrado desde 1 hasta n en la columna denominada “Nº de fallo”. Indica también en la columna “Nº de caso de prueba” el caso de prueba que reveló el fallo.

Conclusiones

11. Recoge el formulario de conclusiones para las pruebas estructurales (E25)
12. Una vez hayas rellenado el formulario de conclusiones, entrega todo el material a la persona a cargo del ejercicio. Has terminado.

Criterio de Cobertura

Obtener cobertura de decisiones (y de sentencias) significa que cada rama en un componente se ha ejecutado al menos una vez. Las ramas en los programas escritos en C las crean las sentencias: `if`, `?`, `for`, `while` y `do-while`. Cada una de estas sentencias genera dos decisiones (ramas); la evaluación de la expresión de la decisión debe tomar valores verdadero y falso al menos una vez.

Las ramas también las puede crear la construcción `switch-case`. Para probar todas las ramas, se deben ejecutar todas las etiquetas `case` al menos una vez. Esto incluye la etiqueta `default`, incluso si no está escrita explícitamente en el código fuente. Cada etiqueta `case` genera una decisión única.

F.14 Formulario de Recogida de Datos para las Pruebas Estructurales (E21)

Nombre: _____

Aspectos de tiempo

Para cada entrada referida al tiempo que ha transcurrido, deduce el tiempo que hayas tomado para descansos, etc.

1. ¿A qué hora comenzaste el ejercicio de prueba estructural? (hora:minutos)
2. ¿Cuánto tiempo necesitaste para elaborar los casos de prueba? (horas:minutos)
3. ¿Cuántos casos de prueba generaste?
4. ¿Cuánto tiempo necesitaste para teclear y ejecutar los casos de prueba? (horas:minutos)
5. ¿A qué hora comenzaste a buscar los fallos? (hora:minutos)
6. ¿Cuánto tiempo necesitaste para identificar los fallos? (horas:minutos)
7. ¿A qué hora terminaste el ejercicio? (hora:minutos)
8. ¿Podrías asegurar que encontraste todas los fallos? Estima el porcentaje de fallos que has encontrado (en %)
9. ¿Cómo de bien crees que has efectuado la prueba estructural? Marca la escala apropiadamente.

Estimación	0	1	2	3	4	5
Comparación	fatal	bastante mal	regular	bien	muy bien	perfectamente

F.15 Formulario de Datos de Prueba para las Pruebas Estructurales (E22)

Nombre: _____

Nº de caso	Propósito del caso	Datos de prueba

Nº de caso	Salida Observada	Salida Esperada

Fallos detectados				
Nº de fallo.	Nº de caso de prueba	Breve descripción del fallo	Seguridad de que es un fallo (%)	Encontrado con la técnica (si/no)

F.18 Formulario de Conclusiones para las Pruebas Estructurales (E25)

Nombre: _____

Sobre la preparación del ejercicio

1. ¿Cuánto tiempo has dedicado a estudiar la técnica funcional? (Indica el número de horas)
2. ¿Cuántos ejercicios has hecho por tu cuenta (fuera de clase) para la técnica funcional? Cuantifícalo con un número incluyendo los ejercicios obligatorios si los hubo.

Sobre la generación de los casos de prueba

3. Indica y numera los pasos que has seguido para aplicar la técnica estructural.
4. ¿Cuánto de seguro estás de que los casos de prueba que has generado ejercitan todas las faltas?

Muy seguro	Medianamente seguro	Poco seguro	Nada seguro
------------	---------------------	-------------	-------------

Sobre la ejecución de los casos de prueba

5. ¿Cuánto de seguro estás de que has detectado todos los fallos que se han revelado durante la ejecución de los casos de prueba?

Muy seguro	Medianamente seguro	Poco seguro	Nada seguro
------------	---------------------	-------------	-------------

6. ¿Has encontrado alguna situación en la que no hayas sido capaz de determinar si estabas frente a un fallo o no? Si En caso afirmativo, explica la situación.

F.19 Código Fuente *cmdline* (E00)

```

1  #ifndef CMDLINEH
2  #define CMDLINEH

4  #define AYUDA  0
5  #define MEDIDA 1
6  #define GKOM   2
7  #define LKOM   3
8  #define LKHM   4
9  #define GKHM   5
10 #define LIHE    6
11 #define GIHE    7
12 #define TODAS   8
13 #define MAX      9
14 #define PROMEDIO 10
15 #define MIN     11
16 #define INF     12
17 #define SUP     13

19 struct entrada_opcion {
20     char *opcion;
21     int long_min;
22     int identificador;
23 };

25 struct comando {
26     double numero;
27     char *medida;
28     char *busqueda;
29 };

31 #endif

```

Listado F.1: cmdline.h

```

1  #include <stdio.h>
2  #include <string.h>
3  #include <ctype.h>
4  #include <stdlib.h>
5  #include "cmdline.h"

7  static struct entrada_opcion tabla_opciones[] = {
8      {"-?", 2, AYUDA},
9      {"-ayuda", 2, AYUDA},
10     {"-medida", 3, MEDIDA},
11     {"GKOM", 3, GKOM},
12     {"LKOM", 3, LKOM},
13     {"LKHM", 3, LKHM},

```

```
14 {"GKHM", 3, GKHM},
15 {"LIHE", 2, LIHE},
16 {"-todas", 2, TODAS},
17 {"-maximo", 3, MAX},
18 {"-promedio", 3, PROMEDIO},
19 {"-minimo", 3, MIN},
20 {"-inferior", 3, INF},
21 {"-superior", 3, SUP},
22 };
23 static int tamano_tabla_opciones = sizeof(tabla_opciones) /
24                                     sizeof(struct entrada_opcion);

27 void uso(char *progn)
28 {
29     fprintf(stderr,
30         "Uso: %s -medida <MEDIDA> [ Opcion-busqueda ] Fichero [ Fichero ... ]\n", progn);
31     exit(1);
32 }

34 static int cadena_es_opcion(char *arg,
35                             char *arg_valido,
36                             int long_min_req)
37 {
38     int long_arg, res;

40     res = 0;
41     long_arg = strlen(arg);
42     if ((strcmp(arg, arg_valido) == 0) && (long_arg >= long_min_req))
43         res = 1;

45     return res;
46 }

49 static int codigo_cadena(char *cadena)
50 {
51     int i, res;

53     res = -1;
54     for (i = 0; i < tamano_tabla_opciones; ++i) {
55         if (cadena_es_opcion(cadena, tabla_opciones[i].opcion, tabla_opciones[i].long_min)) {
56             res = tabla_opciones[i].identificador;
57             break;
58         }
59     }
```

```
60     return res;
61 }

64 void imprimir_comando(struct comando *comando)
65 {
66     if ( (comando->medida != NULL) && (strcmp(comando->medida, "-ayuda") ==
67         0))
68         uso("cmdline");
69     else if ( (comando->medida == NULL) && (comando->busqueda == NULL) )
70         fprintf(stderr, "No se ha proporcionado opcion de medida\n");
71     else {
72         printf("Resumen del comando:\n");
73         if (comando->medida)
74             printf("Medida: %s\n", comando->medida);
75         if (comando->busqueda) {
76             printf("Busqueda: %s\n", comando->busqueda);
77             if (strcmp(comando->busqueda, "-inferior") == 0 || strcmp(comando->
78                 busqueda, "-superior") == 0)
79                 printf("Numero: %.1f\n", comando->numero);
80         }
81     }

83 void imprimir_ficherosin(int argi,
84                         int argc,
85                         char **argv)
86 {
87     int i;

88     if (argi == 0)
89         printf("No se han proporcionado ficheros de entrada\n");
90     else if (argi > 0)
91     {
92         printf("Numero de ficheros de entrada: %d\n", argc - argi);
93         printf("Los ficheros de entrada son:\n");
94         for (i = argi ; i < argc; ++i)
95             printf(" %s\n", argv[i]);
96     }
97 }

101 int procesar_opciones(int argc,
102                     char **argv,
103                     struct comando *comando)
104 {
105     int i, res = 0, posicion = 0;
```

```
107 comando->numero = 0.0;
108 comando->medida = NULL;
109 comando->busqueda = NULL;

111 for (i = 1; i < argc; ++i) {
112     switch (codigo_cadena(argv[i])) {

114         case AYUDA:
115             comando->medida = "-ayuda";
116             posicion = -1;
117             break;

119         case MEDIDA:
120             if (comando->medida != NULL) {
121                 fprintf(stderr, "Demasiadas opciones de medida\n");
122                 res = -1;
123             }
124             else {
125                 if (++i >= argc) {
126                     fprintf(stderr, "No hay busqueda\n");
127                     res = -1;
128                 }
129                 else {
130                     switch (codigo_cadena(argv[i])) {
131                         case GKOM:
132                             comando->medida = "GKOM";
133                             break;
134                         case LKOM:
135                             comando->medida = "LKOM";
136                             break;
137                         case LKHM:
138                             comando->medida = "LKHM";
139                             break;
140                         case GKHM:
141                             comando->medida = "GKHM";
142                             break;
143                         case LIHE:
144                             comando->medida = "LIHE";
145                             break;
146                         default:
147                             fprintf(stderr, "Medida no valida\n");
148                             res = -1;
149                             break;
150                     }
151                 }
152             }
153             break;
```

```

155     case TODAS:
156         if (comando->busqueda != NULL) {
157             fprintf(stderr, "Demasiadas opciones de busqueda\n");
158             res = -1;
159         }
160         else
161             comando->busqueda = "-todas";
162         break;

164     case MAX:
165         if (comando->busqueda != NULL) {
166             fprintf(stderr, "Demasiadas opciones de busqueda\n");
167             res = -1;
168         }
169         else
170             comando->busqueda = "-maximo";
171         break;

173     case PROMEDIO:
174         if (comando->busqueda != NULL) {
175             fprintf(stderr, "Demasiadas opciones de busqueda\n");
176             res = -1;
177         }
178         else
179             comando->busqueda = "-promedio";
180         break;

182     case MIN:
183         if (comando->busqueda != NULL) {
184             fprintf(stderr, "Demasiadas opciones de busqueda\n");
185             res = -1;
186         }
187         else
188             comando->busqueda = "-minimo";
189         break;
190     case INF:
191         if (comando->busqueda != NULL) {
192             res = -1;
193         }
194         else {
195             comando->busqueda = "-inferior";
196             ++i;
197             if (i >= argc ||
198                 !(isdigit(*argv[i]) || (*argv[i]=='-' && isdigit(*(argv[i]+sizeof
199                     (char)))))) {
200                 fprintf(stderr, "No se ha proporcionado un numero\n");
201                 res = -1;

```

```
201 }
202 else {
203     comando->numero = atoi(argv[i]);
204 }
205 }
206 break;

208 case SUP:
209     if (comando->busqueda != NULL) {
210         fprintf(stderr, "Demasiadas opciones de busqueda\n");
211         res = -1;
212     }
213     else {
214         comando->busqueda = "-superior";
215         ++i;
216         if (i >= argc ||
217             !(isdigit(*argv[i]) || (*argv[i]=='-' && isdigit(*(argv[i]+sizeof
218                 (char)))))) {
219             fprintf(stderr, "No se ha proporcionado un numero\n");
220             res = -1;
221         }
222         else {
223             comando->numero = atoi(argv[i]);
224         }
225         break;

227     default:
228         posicion = i;
229         break;
230     }

232     if ( (res < 0) || (posicion == i) )
233         break;
234 }

236 if (res < 0)
237     fprintf(stderr, "Los argumentos de la opcion son incorrectos\n");
238 else if (comando -> medida == NULL)
239     imprimir_comando(comando);
240 else if (posicion == 0)
241     imprimir_ficherosin(posicion, argc, argv);
242 else {
243     imprimir_comando(comando);
244     imprimir_ficherosin(posicion, argc, argv);
245 }
246 return res;
247 }
```

```
250 /*****
251  * Este es el programa principal.
252  * No hay que crear datos de prueba para este, crear abstracciones, etc.
253  *****/
254 int main(int argc,
255          char **argv)
256 {
257     struct comando a;
258     int res;

260     if (argc == 1) {
261         uso("cmdline");
262         res = -1;
263     }
264     else
265         res = procesar_opciones(argc, argv, &a);

267     return res;
268 }
```

Listado F.2: cmdline.c

F.20 Especificación del programa “cmdline” (E01)

Nombre

cmdline – Analiza sintáctica y semánticamente una línea de comandos

Uso

cmdline -medida <MEDIDA> [Opción-búsqueda] Fichero [Fichero...]

Descripción

cmdline analiza la línea de comandos proporcionada y evalúa su corrección sintáctica y semántica. El usuario debe proporcionar como argumentos obligatorios al programa una medida y al menos un fichero; además, se puede proporcionar como argumento opcional una opción de búsqueda (que ha de ser única). El orden en el que aparecen la opción de medida y la opción de búsqueda no es importante. Cuando la línea de comandos se evalúa, todos los argumentos que aparecen a partir del primero que no sea reconocido como una opción se tratan como nombres de ficheros. La existencia de tales ficheros no se comprueba.

En caso de éxito (es decir, la sintaxis y el significado de los argumentos es válido), se presenta por pantalla un resumen de los argumentos. El resumen consiste en la medida que ha de ser computada, las opciones de búsqueda en caso de que las hubiese (incluyendo el valor para las opciones **-inferior** y **-superior**), y una lista de los ficheros que han de ser leídos.

En otro caso, se muestra un mensaje de error, que variará dependiendo del mismo. Los mensajes que pueden aparecer se refieren a:

- Ausencia de los ficheros de entrada. El error que aparecerá por pantalla será: “No se han proporcionado ficheros de entrada”.
- Ausencia de la opción de medida. El error que aparecerá por pantalla será: “No se ha proporcionado opcion de medida”.
- Opciones incorrectas (medida u opción de búsqueda). Aparecerá por pantalla el error: “Los argumentos de la opcion son incorrectos”, junto con alguno de los siguientes mensajes para aclarar el problema:
 - Si se proporciona más de varias veces la opción **-medida** aparecerá el error: “Demasiadas opciones de medida”.
 - Si no se proporciona opción de medida, aparecerá el error: “No hay opcion de medida”.
 - Si se proporciona una opción de medida no válida, aparecerá el error: “Medida no valida”.
 - Si se proporcionan varias opciones de búsqueda aparecerá el error: “Demasiadas opciones de busqueda”.
 - Si no se proporciona número en las opciones de búsqueda que lo precisen, aparecerá el error: “No se ha proporcionado un numero”.

Opciones

Las opciones y los argumentos para las opciones se pueden escribir completos o reducirse a formas cortas que corresponden a sus prefijos. A continuación se muestran tanto las opciones y sus argumentos como los prefijos usados para cada una de las opciones (el prefijo corresponde a lo que hay antes de los corchetes, y la opción completa al prefijo más lo que hay entre corchetes). Se puede usar como máximo una opción de búsqueda y otra de medida.

- **-a[yuda]**
Opción de ayuda. Se saca por pantalla un texto acerca del uso y no se hace nada más.
- **-?**
Opción de ayuda. Funciona igual que la anterior.
- **-me[dida] <MEDIDA>**
Las medidas aceptadas son GKO[M], LKO[M], GKH[M], LKH[M], GI[HE] y LI[HE].
- **-t[odas]**
Opción de búsqueda. La medida debe ser evaluada para todas las unidades.
- **-ma[ximo]**
Opción de búsqueda. Determina el valor máximo de la medida para todas las unidades.
- **-mi[nimo]**
Opción de búsqueda. Determina el valor mínimo de la medida para todas las unidades.
- **-pr[omedio]**
Opción de búsqueda. Determina el valor medio de la medida para todas las unidades.
- **-in[ferior] <valor>**
Opción de búsqueda. Todas las unidades para las cuales la medida cae bajo el límite serán identificadas. El <valor> debe ser un número real.
- **-su[perior] <valor>**
Opción de búsqueda. Todas las unidades para las cuales la medida supera el límite serán identificadas. El <valor> debe ser un número real.

F.21 Suplemento “Revisión de código” y “cmdline” (E02)

Ésta es la hoja suplementaria para aplicar la técnica “revisión de código” al programa `cmdline`.

Visión general

El programa `cmdline` realiza simplemente un parsing de la línea de comandos para determinar si es válida. En dicho caso, muestra una descripción de la línea de comandos introducida o una indicación de por qué es incorrecta. Es importante notar que no lleva a cabo ninguno de los comandos, sino que únicamente valida la entrada.

A continuación aparece un ejemplo sencillo de su uso:

```
C:\ cmdline -medida GKOM fichero
Resumen del comando:
Medida: GKOM
Numero de ficheros de entrada: 1
Los ficheros de entrada son:
fichero

C:\ cmdline -medida JUNK fichero
Opcion de medida no valida
Los argumentos de la opcion son incorrectos
```

La existencia de los ficheros de entrada no se comprueba, por lo que no tienen que ser creados. Las medidas (`GKOM` en el ejemplo anterior) son acrónimos de operaciones que deben llevarse a cabo sobre los datos de los ficheros (para el propósito de este ejercicio no es necesario que se sepa lo que son).

Breve descripción de las funciones de librería usadas

- `int strlen(char *s)`
Devuelve la longitud de la cadena `s`
- `int strcmp(char *s1, char *s2)`
Compara dos cadenas. Devuelve un valor mayor que, igual o menor que cero dependiendo de si `s1` es lexicográficamente (en valor ASCII) mayor que, igual o menor que `s2`.
- `int strncmp(char *s1, char *s2, int n)`
Igual que la anterior, pero comparando un máximo de `n` caracteres.
- `int fprintf(stderr, "Texto"...)`
Imprime el texto en la salida error estándar.
- `int isdigit(char c)`
Devuelve *verdadero* si el carácter es un dígito, *falso* en otro caso.
- `int atoi(char *str)`
Convierte la cadena `str` en un entero. Los espacios iniciales se ignoran, y la transformación de la cadena termina con el primer carácter no numérico. Si no se pudieron leer caracteres no numéricos, se devuelve 0.

Recordatorio

No se deben generar abstracciones para la función principal del programa.

F.22 Suplemento “Pruebas funcionales” y “cmdline” (E03)

Ésta es la hoja suplementaria para aplicar la técnica “pruebas funcionales” al programa `cmdline`.

Visión general

El programa `cmdline` realiza simplemente un parsing de la línea de comandos para determinar si es válida. En dicho caso, muestra una descripción de la línea de comandos introducida o una indicación de por qué es incorrecta. Es importante notar que no lleva a cabo ninguno de los comandos, sino que únicamente valida la entrada.

A continuación aparece un ejemplo sencillo de su uso:

```
C:\> cmdline -medida GKOM fichero
Resumen del comando:
Medida: GKOM
Numero de ficheros de entrada: 1
Los ficheros de entrada son:
fichero

C:\> cmdline -medida JUNK fichero
Opcion de medida no valida
Los argumentos de la opcion son incorrectos
```

La existencia de los ficheros de entrada no se comprueba, por lo que no tienen que ser creados. Las medidas (GKOM en el ejemplo anterior) son acrónimos de operaciones que deben llevarse a cabo sobre los datos de los ficheros (para el propósito de este ejercicio no es necesario que se sepa lo que son).

Documentos y código usados

Para este ejercicio es necesario el documento E01, que incluye la especificación del componente, así como el código fuente del programa, que será entregado después de que haber escrito los casos de prueba.

Búsqueda de los ficheros necesarios

Crea un nuevo directorio para evitar conflictos con cualquier otro ejercicio de pruebas y copia el fichero `cmdline.exe` del directorio que se ha indicado.

Creación de clases de equivalencia y casos de prueba

El componente “`cmdline`” se puede probar directamente. No son necesarios ficheros de entrada.

F.23 Suplemento “Pruebas estructurales” y “cmdline” (E04)

Ésta es la hoja suplementaria para aplicar la técnica “pruebas estructurales” al programa `cmdline`.

Visión general

El programa `cmdline` realiza simplemente un parsing de la línea de comandos para determinar si es válida. En dicho caso, muestra una descripción de la línea de comandos introducida o una indicación de por qué es incorrecta. Es importante notar que no lleva a cabo ninguno de los comandos, sino que únicamente valida la entrada.

A continuación aparece un ejemplo sencillo de su uso:

```
C:\ cmdline -medida GKOM fichero
Resumen del comando:
Medida: GKOM
Numero de ficheros de entrada: 1
Los ficheros de entrada son:
fichero

C:\ cmdline -medida JUNK fichero
Opcion de medida no valida
Los argumentos de la opcion son incorrectos
```

La existencia de los ficheros de entrada no se comprueba, por lo que no tienen que ser creados. Las medidas (GKOM en el ejemplo anterior) son acrónimos de operaciones que deben llevarse a cabo sobre los datos de los ficheros (para el propósito de este ejercicio no es necesario que se sepa lo que son).

Documentos y código usados

Para este ejercicio es necesario el código fuente del programa así como el documento E01, que contiene la especificación del componente, y que será entregado después de haber creados los casos de prueba y haber intentado conseguir la cobertura pedida.

Búsqueda de los ficheros necesarios

Crea un nuevo directorio para evitar conflictos con cualquier otro ejercicio de pruebas y copia el fichero `cmdline.exe` del directorio indicado.

Creación de los casos de prueba

Todas las funciones de programa se prueban vía la invocación:

```
cmdline [ -argumento ... ]
```

“`cmdline`” no necesita que se creen ficheros de entrada.

Los argumentos que se permiten en la línea de comandos y las funciones que invocan dichos comandos se pueden ver en la función `procesar_opciones` en el fichero `cmdline.c`. La función principal no necesita ser probada.

Breve descripción de las funciones de librería usadas

- `int strlen(char *s)`
Devuelve la longitud de la cadena `s`

- `int strcmp(char *s1, char *s2)`

Compara dos cadenas. Devuelve un valor mayor que, igual o menor que cero dependiendo de si `s1` es lexicográficamente (en valor ASCII) mayor que, igual o menor que `s2`.

- `int strncmp(char *s1, char *s2, int n)`

Igual que la anterior, pero comparando un máximo de `n` caracteres.

- `int fprintf(stderr, "Texto"...)`

Imprime el texto en la salida error estándar.

- `int isdigit(char c)`

Devuelve *verdadero* si el carácter es un dígito, *falso* en otro caso.

- `int atoi(char *str)`

Convierte la cadena `str` en un entero. Los espacios iniciales se ignoran, y la transformación de la cadena termina con el primer carácter no numérico. Si no se pudieron leer caracteres no numéricos, se devuelve 0.

F.24 Código Fuente *nametbl* (E00)

```
1  #ifndef NAMETBLH
2  #define NAMETBLH

4  enum tipo_objeto {
5      SISTEMA,
6      RECURSO,
7      NO_INFO};

9  static char *Nombre_tipo_objeto[] = {
10     "SISTEMA",
11     "RECURSO",
12     "NO_INFO"};

14  enum tipo_recurso {
15      R_SISTEMA,
16      FUNCION,
17      DATO,
18      MEMORIA,
19      R_NO_INFO};

21  static char *Nombre_tipo_recurso[] = {
22     "R_SISTEMA",
23     "FUNCION",
24     "DATO",
25     "MEMORIA",
26     "R_NO_INFO"};

29  struct elemento_tabla {
30      char *nombre;
31      enum tipo_objeto objeto;
32      enum tipo_recurso recurso;
33  };

36  typedef struct elemento_tabla ELEMENTO;

38  struct tabla {
39      char *raiz;
40      int n_elementos;
41  };

43  typedef struct tabla TABLA;

45  #endif
```

Listado F.3: *nametbl.h*

```
1  #include <stdio.h>
2  #include <assert.h>
3  #include <string.h>
4  #include <search.h>
5  #include "nametbl.h"
```

```
7 TABLA *crear_tabla(void)
8 {
9     TABLA *ptr;

11     ptr = (TABLA *)malloc(sizeof(TABLA));
12     assert(ptr != NULL);
13     ptr->raiz = NULL;
14     ptr->n_elementos = 0;

16     return ptr;
17 }

20 int cuantos(TABLA *tabla)
21 {
22     return tabla->n_elementos;
23 }

25 int comparar_elementos(ELEMENTO *p1,
26                         ELEMENTO *p2)
27 {
28     int longitud1, longitud2;

30     longitud1 = strlen(p1->nombre);
31     longitud2 = strlen(p2->nombre);

33     if (longitud1 < longitud2)
34         return(strncmp(p1->nombre, p2->nombre, longitud1));
35     else return(strncmp(p1->nombre, p2->nombre, longitud2));
36 }

39 void imprimir_elemento(ELEMENTO *nodo)
40 {
41     char nombre[7]="Nombre";
42     char tobjeto[9]="T_Objeto";
43     char trecurso[10]="T_Recurso";

45     printf("%s : %s\n", nombre, nodo->nombre);
46     printf("%s : %s\n", tobjeto, Nombre_tipo_objeto[nodo->objeto]);
47     printf("%s : %s\n", trecurso, Nombre_tipo_recurso[nodo->recurso]);
48     printf("-----\n");
49 }

52 void imprimir_en_ultima_visita(ELEMENTO **nodop,
53                                VISIT orden,
54                                int nivel)
55 {
56     if (orden == postorder || orden == leaf)
57         (void) imprimir_elemento(*nodop);
58 }
```



```
61 void insertar_elemento(TABLA *tabla,
62                        char *nombre,
63                        enum tipo_objeto objeto,
64                        enum tipo_recurso recurso)
65 {
66     ELEMENTO *elemento, **resultado;
67
68     elemento = (ELEMENTO *) malloc(sizeof(ELEMENTO));
69     assert(elemento != NULL);
70
71     elemento->nombre = strdup(nombre);
72     assert(elemento->nombre != NULL);
73     elemento->objeto = objeto;
74     elemento->recurso = recurso;
75     (tabla->n_elementos)++;
76
77     resultado = (ELEMENTO **) tsearch((char *)elemento, &(tabla->raiz),
78                                       comparar_elementos);
79     assert((char *) *resultado == (char *) elemento);
80 }
81
82 ELEMENTO *recuperar_elemento(TABLA *tabla,
83                              char *string)
84 {
85     ELEMENTO **elemento, *resultado, clave;
86
87     clave.nombre = string;
88     elemento = (ELEMENTO **) tfind( (char *)&clave, &(tabla->raiz),
89                                     comparar_elementos);
90     if (elemento != NULL)
91         resultado = *elemento;
92     else
93         resultado = NULL;
94
95     return resultado;
96 }
97
98
99
100 int mi_asignar_objeto (TABLA *tabla,
101                       char *nombre,
102                       enum tipo_objeto objeto)
103 {
104     int resultado;
105     ELEMENTO *elemento;
106
107     resultado = -1;
108     elemento = recuperar_elemento(tabla, nombre);
109     if (elemento != NULL) {
110         resultado = 0;
111         elemento->objeto = objeto;
112     }
```

```

114     return resultado;
115 }

118 int mi_asignar_recurso (TABLA *tabla,
119                        char *nombre,
120                        enum tipo_recurso recurso)
121 {
122     int resultado;
123     ELEMENTO *elemento;

125     resultado = -1;
126     elemento = recuperar_elemento(tabla, nombre);
127     if (elemento != NULL) {
128         resultado = 0;
129         elemento->recurso = recurso;
130     }

132     return resultado;
133 }

135 int insertar_simbolo(TABLA *tabla,
136                     char *nombre)
137 {
138     int resultado=0;

140     insertar_elemento(tabla, nombre, NO_INFO, R_NO_INFO);

142     return resultado;
143 }

145 int asignar_objeto(TABLA *tabla,
146                  char *nombre,
147                  char *str_objeto)
148 {
149     int resultado;
150     enum tipo_objeto objeto;

152     objeto = -1;
153     if (strcmp("SISTEMA", str_objeto) == 0)
154         objeto = SISTEMA;
155     else if (strcmp("RECURSO", str_objeto) == 0)
156         objeto = RECURSO;

158     if (objeto == -1) {
159         printf("Tipo de objeto '%s' no reconocido\n", str_objeto);
160         resultado = -1;
161     }
162     else {
163         resultado = mi_asignar_objeto(tabla, nombre, objeto);
164         if (resultado)
165             printf("El nombre '%s' no pertenece a la Tabla.\n", nombre);
166     }

168     return resultado;

```

```
169 }

172 int asignar_recurso(TABLA *tabla,
173                    char *nombre,
174                    char *str_recurso)
175 {
176     int resultado;
177     ELEMENTO *elemento;
178     enum tipo_recurso recurso;

180     recurso = -1;
181     elemento = recuperar_elemento (tabla,nombre);

183     if (strcmp("R_SISTEMA", str_recurso) == 0)
184         recurso = R_SISTEMA;
185     else if (strcmp("FUNCION", str_recurso) == 0)
186         recurso = FUNCION;
187     else if (strcmp("DATO", str_recurso) == 0)
188         recurso = DATO;
189     else if (strcmp("MEMORIA", str_recurso) == 0)
190         recurso = MEMORIA;

192     if ((recurso == -1) && (elemento == NULL)) {
193         printf("La operacion no se ha podido realizar.\n");
194     } else if (recurso == -1) {
195         printf("Tipo de recurso '%s' no reconocido\n", str_recurso);
196         resultado = -1;
197     }
198     else {
199         resultado = mi_asignar_recurso(tabla, nombre, recurso);
200         if (resultado)
201             printf("El nombre '%s' no pertenece a la Tabla.\n", nombre);
202     }

204     return resultado;
205 }

208 void buscar_simbolo(TABLA *tabla,
209                    char *nombre)
210 {
211     ELEMENTO *elemento;

213     printf("Buscando el Nombre '%s':\n", nombre);
214     elemento = recuperar_elemento(tabla, nombre);
215     if (elemento != NULL)
216         imprimir_elemento(elemento);
217     else printf("El nombre '%s' no pertenece a la Tabla.\n", nombre);
218 }

220 void imprimir_tabla(TABLA *tabla)
221 {
222     printf("La Tabla tiene las siguientes %d entradas:\n", cuantos(tabla));
223     twalk((char *)tabla->raiz, imprimir_en_ultima_visita);
```

```

224 }

227 /*****
228  * De aquí en adelante está el programa para leer del fichero de entrada
229  * No crear casos de prueba ni construir abstracciones
230  * para esta parte del código.
231  *****/

233 void procesar_comandos(filep)
234     FILE *filep;
235 {
236     TABLA *nt;
237     char *ptr;
238     char buf[BUFSIZ], command[BUFSIZ], name[BUFSIZ], typ[BUFSIZ];

240     nt = crear_tabla();
241     assert (nt != NULL);

243     while (fgets(buf, sizeof(buf), filep) != NULL) {
244         if ( (ptr = strchr(buf, '\n')) != NULL)
245             *ptr = '\0';
246         printf("\nProcesando la linea '%s':\n", buf);
247         *command = '\0';
248         *name = '\0';
249         *typ = '\0';
250         sscanf(buf, "%s %s %s", command, name, typ);
251         if (strcmp(command, "insertar_simbolo") == 0)
252             insertar_simbolo(nt, name);
253         else if (strcmp(command, "asignar_objeto") == 0)
254             asignar_objeto(nt, name, typ);
255         else if (strcmp(command, "asignar_recurso") == 0)
256             asignar_recurso(nt, name, typ);
257         else if (strcmp(command, "buscar_simbolo") == 0)
258             buscar_simbolo(nt, name);
259         else if (strcmp(command, "imprimir_tabla") == 0)
260             imprimir_tabla(nt);
261         else printf("Comando '%s' no reconocido\n", command);
262     }
263 }

265 int main(argc, argv)
266     int argc; char **argv;
267 {
268     TABLA *nt;
269     FILE *filep;
270     char *file;

272     if (argc != 2)
273         fprintf(stderr, "Uso: %s Dato\n", *argv);
274     else {
275         file = argv[1];
276         if ((filep = fopen(file, "r")) == NULL)
277             perror(file);
278         else {

```

```
279     printf("Procesando el fichero de entrada '%s'.\n", file);
280     procesar_comandos(filep, nt);
281     printf("Fin de fichero de entrada '%s'.\n", file);
282     fclose(filep);
283 }
284 }
285 return 0;
286 }
```

Listado F.4: nametbl.c

F.25 Especificación del programa “nametbl” (E01)

Nombre

nametbl – Funciones para manejar una tabla de símbolos

Uso

nametbl <fichero-de-entrada>

Descripción

nametbl lee comandos de un fichero y los procesa para probar una serie de funciones, las cuales permiten manejar el tipo abstracto de datos correspondiente a una tabla de símbolos para un determinado lenguaje. La tabla de símbolos almacena la siguiente información para cada uno de los símbolos:

- Nombre del símbolo.
- Tipo de objeto del símbolo: {NO_INFO, SISTEMA, RECURSO, COMBINADO}
- Tipo de recurso del símbolo: {R_NO_INFO, R_SISTEMA, FUNCION, DATO}

Opciones

No soporta ninguna opción

Datos de Entrada

Sólo se permiten los siguientes comandos en el fichero de entrada. Cada comando debe comenzar en una nueva línea. El uso de mayúsculas y minúsculas se distingue tanto para los símbolos como para los identificadores de objeto y recurso. El separador entre comandos y argumentos es el espacio; lo que significa que no puede haber espacios dentro de una cadena.

- **insertar_simbolo** <símbolo>
Inserta el símbolo en la tabla. Se asigna como tipo de objeto del símbolo el valor **NO_INFO** y como tipo de recurso el valor **R_NO_INFO**. En caso de que se intente insertar el mismo símbolo dos veces, aparecerá un mensaje de error.
- **asignar_objeto** <símbolo> <tipo-objeto>
Asigna el tipo de objeto <tipo-objeto> al símbolo <símbolo>, donde <tipo-objeto> debe ser **SISTEMA** o **RECURSO** o **COMBINADO**. Cualquier otro tipo de objeto generará un mensaje de error, así como el valor **NO_INFO**, que no se permite. El símbolo debe pertenecer a la tabla, o si no se generará el correspondiente mensaje de error. El valor anterior que tuviese el tipo de objeto se sobrescribe y por tanto se pierde.
- **asignar_recurso** <símbolo> <tipo-recurso>
Asigna el tipo de recurso <tipo-recurso> al símbolo <símbolo>, donde <tipo-recurso> debe ser **R_SISTEMA** o **FUNCION** o **DATO**. Cualquier otro tipo de recurso generará un mensaje de error, así como el valor **R_NO_INFO**, que no se permite. El valor **R_SISTEMA** sólo podrá usarse en el caso en que el tipo de objeto del símbolo tiene el valor **SISTEMA**, en caso contrario, se generará un mensaje de error. El símbolo debe pertenecer a la tabla, o si no se generará el correspondiente mensaje de error. El valor anterior que tuviese el tipo de recurso se sobrescribe y por tanto se pierde.
- **buscat_simbolo** <símbolo>
Busca el símbolo <símbolo> e imprime la correspondiente información. Si no se encuentra el símbolo en la tabla, se genera el correspondiente mensaje de error.

- `imprimir_tabla`

Imprime el número de entradas de la tabla de símbolos así como su contenido.

Limitaciones

La funcionalidad para leer ficheros es para propósito de las pruebas y de este modo no es especialmente tolerante a fallos. Por ejemplo, uno puede asumir que comandos que están escritos de forma incorrecta, comandos con argumentos que faltan, o comandos con demasiados argumentos no se comprobarán con demasiado detalle.

F.26 Suplemento “Revisión de código” y “nametbl” (E02)

Ésta es la hoja suplementaria para aplicar la técnica “revisión de código” al programa `nametbl`.

Visión general

El programa proporciona funciones para construir una tabla de símbolos sencillas y para asignar atributos a los símbolos. A continuación aparece un ejemplo sencillo de su uso:

```
C:\ type entrada
insertar_simbolo qwe
insertar_simbolo rty
imprimir_tabla

C:\ nametbl entrada
Procesando el fichero de entrada 'entrada'.
Procesando la linea insertar_simbolo qwe':
Procesando la linea 'insertar_simbolo rty':
Procesando la linea 'imprimir_tabla':
La Tabla tiene las siguientes 2 entradas:
Nombre: qwe
Objeto: NO_INFO
Recurso: R_NO_INFO
-----
Nombre: rty
Objeto: NO_INFO
Recurso: R_NO_INFO
-----
Fin del fichero de entrada 'entrada'
```

Breve descripción de las funciones de librería usadas

- `assert(expresión)`
Macro que espera que la expresión que se le proporciona sea cierta en tiempo de ejecución, y en ese caso no hace nada. En otro caso genera un mensaje de error mostrando la expresión, y el programa termina.
- `char* malloc(unsigned tamaño)`
Devuelve un puntero a un bloque de memoria que contiene al menos `tamaño` bytes. En caso de fallo, devuelve `NULL`.
- `int free(ptr)`
Devuelve un bloque de memoria pedido anteriormente al sistema. `ptr` debe haber sido pedido usando `malloc` o una función similar. En caso de éxito, devuelve 1, y si no devuelve 0.
- `int strcmp(s1,s2)`
Compara dos cadenas. Devuelve un valor mayor, igual o menor que 0 dependiendo de si `s1` es lexicográficamente hablando mayor, igual o menor que `s2`.
- `char *strdup(s1)`
Duplica la cadena `s1`. Reserva memoria para este propósito usando `malloc()` y devuelve un puntero a la cadena duplicada en caso de éxito y si no, devuelve `NULL`.

- `int sscanf(s, formato, punteros)`

Lee caracteres de la cadena `s`, los convierte al formato indicado y escribe el resultado en las variables a las que apunta `punteros`. Es la función contraria a `printf`; las cadenas de especificación de formato son las mismas.

- `char *fgets(s, n, flujo)`

Lee caracteres de `flujo` y los escribe en la cadena `s`, hasta que se hayan leído `n-1` caracteres o se haya encontrado un salto de línea o un fin de fichero.

- `char *tsearch(char *clave, char **ptr_raiz, int (*compar)())`
- `char *tfind(char *clave, char **ptr_raiz, int (*compar)())`
- `void twalk(char *raiz, void (*accion)())`

Estas funciones implementan un árbol de búsqueda binario. Todas las comparaciones requeridas las realiza la función `compar`, que debe ser proporcionada por el usuario para que usen las funciones del árbol. La función `compar` se suministra a las funciones del árbol pasándola como un puntero a una función. `compar` se invoca con dos argumentos, que apuntan a los elementos a ser comparados. El valor que devuelve debe ser menor, igual o mayor que cero, dependiendo de si el primer elemento es menor, igual o mayor que el segundo (análogo a `strcmp`).

`tsearch` puede construir y buscar en un árbol. Si se encuentra un elemento con la misma clave que elemento al cual `clave` apunta, se devuelve un puntero que apunta a un puntero que a su vez apunta al elemento que se encontró. Si no lo encuentra, el elemento se inserta en el árbol y se devuelve un puntero a un puntero al nuevo elemento. A continuación se muestra una ilustración representando la doble indirección:

```
puntero    +--+          +-----+
----->|  |----->|  nodo  |
devuelto   +--+ puntero  +-----+
                        /  \
```

Sólo se copian los punteros; es decir, el usuario es responsable de reservar memoria para los tres elementos (por ejemplo con `malloc`), `ptr_raiz` apunta a una variable y esta variable apunta a la raíz del árbol. Para generar el primer nodo del árbol (la raíz), se debe asignar `NULL` a la variable; de este modo la variable recibe un puntero a la raíz del árbol.

`tfind` es idéntica a `tsearch`, con la excepción de que si el elemento no se encuentra, se devuelve `NULL` y no se añade ningún nodo al árbol.

`twalk` recorre un árbol o un subárbol, ya que se puede usar cualquier nodo como argumento raíz. Cabe notar que se da un puntero a `raiz`, no un puntero a un puntero como usan `tsearch` y `tfind`. `accion` es la función que se invoca para cada nodo del árbol. Esta función debe ser proporcionada también por el usuario. Recibe tres argumentos: un puntero al elemento del árbol, información acerca de cuándo se visita el elemento, y la profundidad del elemento del árbol (`raiz=0`).

La información de la visita es un tipo enumerado `typedef enum preorder, postorder, endorder, leaf VISIT`; donde `preorder`, `postorder` y `endorder` representan la primera, segunda y tercera visitas durante una búsqueda en profundidad de izquierda a derecha del árbol. `leaf` se refiere a la hoja de un árbol.

Recordatorio

No se deben generar abstracciones para la función principal del programa.

F.27 Suplemento “Pruebas funcionales” y “nametbl” (E03)

Ésta es la hoja suplementaria para aplicar la técnica “pruebas funcionales” al programa nametbl.

Visión general

El programa proporciona funciones para construir una tabla de símbolos sencilla y para asignar atributos a los símbolos. A continuación aparece un ejemplo sencillo de su uso:

```
C:\ type entrada
insertar_simbolo qwe
insertar_simbolo rty
imprimir_tabla

C:\ nametbl entrada
Procesando el fichero de entrada 'entrada'.
Procesando la linea insertar_simbolo qwe':
Procesando la linea 'insertar_simbolo rty':
Procesando la linea 'imprimir_tabla':
La Tabla tiene las siguientes 2 entradas:
Nombre: qwe
Objeto: NO_INFO
Recurso: R_NO_INFO
-----
Nombre: rty
Objeto: NO_INFO
Recurso: R_NO_INFO
-----
Fin del fichero de entrada 'entrada'
```

Documentos y códigos usados

Para este ejercicio es necesario el documento E01, que incluye la especificación del componente, así como el código fuente del programa, que será entregado después de que haber escrito los casos de prueba.

Búsqueda de los ficheros necesarios

Crea un nuevo directorio para evitar conflictos con cualquier otro ejercicio de pruebas y copia el fichero nametbl.exe del directorio que se ha indicado.

Creación de clases de equivalencia y casos de prueba

Este programa contiene una función para procesar el fichero de entrada y simplificar el uso del programa, que no es necesario probar. *No crees clases de equivalencia para comandos sintácticamente incorrectos*, ya que eso probaría la función de procesamiento. Ejemplos de casos de prueba no necesarios son:

```
borrar todo (El comando no está definido)
insertar_simbolo (Falta el parámetro)
```

Recordatorio

Asegúrate de que los casos de prueba que elabores generen salidas tales que te permitan detectar fallos.

F.28 Suplemento “Pruebas estructurales” y “nametbl” (E04)

Ésta es la hoja suplementaria para aplicar la técnica “pruebas estructurales” al programa nametbl. Este ejercicio prueba varias funciones escritas en C. Para que no malgastes esfuerzos y tiempo en este ejercicio, se ha creado una función que se encarga de leer y procesar el fichero de entrada. No es necesario que pruebes ésta función.

Visión general

El programa proporciona funciones para construir una tabla de símbolos sencilla y para asignar atributos a los símbolos. A continuación aparece un ejemplo sencillo de su uso:

```
C:\ type entrada
insertar_simbolo qwe
insertar_simbolo rty
imprimir_tabla

C:\ nametbl entrada
Procesando el fichero de entrada 'entrada'.
Procesando la linea insertar_simbolo qwe':
Procesando la linea 'insertar_simbolo rty':
Procesando la linea 'imprimir_tabla':
La Tabla tiene las siguientes 2 entradas:
Nombre: qwe
Objeto: NO_INFO
Recurso: R_NO_INFO
-----
Nombre: rty
Objeto: NO_INFO
Recurso: R_NO_INFO
-----
Fin del fichero de entrada 'entrada'
```

Documentos y códigos usados

Para este ejercicio es necesario el código fuente del programa así como el documento E01, que contiene la especificación del componente, y que será entregado después de haber creados los casos de prueba y haber intentado conseguir la cobertura pedida.

Búsqueda de los ficheros necesarios

Crea un nuevo directorio para evitar conflictos con cualquier otro ejercicio de pruebas y copia el fichero nametbl.exe del directorio indicado.

Creación de los casos de prueba

Todas las funciones de programa se prueban vía la invocación:

```
nametbl <fichero_de_entrada>
```

El formato del fichero de entrada es el siguiente:

```
comando1 parametro1 parametro2
comando2 parametroX
...
```

Los comandos permitidos en la entrada y las funciones que invocan dichos comandos aparecen en la función `procesar_comandos` en el fichero `nametbl.c`. Debido a que esta función se ha creado para facilitar las pruebas y no será parte de la aplicación, no es especialmente robusta. Puedes asumir que la función, aunque es sensible, no tiene faltas. No escribas ningún caso de prueba para ejercitar esta función.

Breve descripción de las funciones de librería usadas

- `assert(expresión)`
Macro que espera que la expresión que se le proporciona sea cierta en tiempo de ejecución, y en ese caso no hace nada. En otro caso genera un mensaje de error mostrando la expresión, y el programa termina.
- `char* malloc(unsigned tamaño)`
Devuelve un puntero a un bloque de memoria que contiene al menos `tamaño` bytes. En caso de fallo, devuelve `NULL`.
- `int free(ptr)`
Devuelve un bloque de memoria pedido anteriormente al sistema. `ptr` debe haber sido pedido usando `malloc` o una función similar. En caso de éxito, devuelve 1, y si no devuelve 0.
- `int strcmp(s1,s2)`
Compara dos cadenas. Devuelve un valor mayor, igual o menor que 0 dependiendo de si `s1` es lexicográficamente hablando mayor, igual o menor que `s2`.
- `char *strdup(s1)`
Duplica la cadena `s1`. Reserva memoria para este propósito usando `malloc()` y devuelve un puntero a la cadena duplicada en caso de éxito y si no, devuelve `NULL`.
- `int sscanf(s,formato,punteros)`
Lee caracteres de la cadena `s`, los convierte al formato indicado y escribe el resultado en las variables a las que apunta `punteros`. Es la función contraria a `printf`; las cadenas de especificación de formato son las mismas.
- `char *fgets(s,n,flujo)`
Lee caracteres de `flujo` y los escribe en la cadena `s`, hasta que se hayan leído `n-1` caracteres o se haya encontrado un salto de línea o un fin de fichero.
- `char *tsearch(char *clave, char **ptr_raiz, int (*compar)())`
- `char *tfind(char *clave, char **ptr_raiz, int (*compar)())`
- `void twalk(char *raiz, void (*accion)())`

Estas funciones implementan un árbol de búsqueda binario. Todas las comparaciones requeridas las realiza la función `compar`, que debe ser proporcionada por el usuario para que usen las funciones del árbol. La función `compar` se suministra a las funciones del árbol pasándola como un puntero a una función. `compar` se invoca con dos argumentos, que apuntan a los elementos a ser comparados. El valor que devuelve debe ser menor, igual o mayor que cero, dependiendo de si el primer elemento es menor, igual o mayor que el segundo (análogo a `strcmp`).

`tsearch` puede construir y buscar en un árbol. Si se encuentra un elemento con la misma clave que elemento al cual `clave` apunta, se devuelve un puntero que apunta a un puntero que a su vez apunta al elemento que se encontró. Si no lo encuentra, el elemento se inserta en el árbol y se devuelve un puntero a un puntero al nuevo elemento. A continuación se muestra una ilustración representando la doble indirección:

```
puntero    +--+          +-----+
----->|  |----->|  nodo  |
devuelto   +--+ puntero  +-----+
                        /  \
```

Sólo se copian los punteros; es decir, el usuario es responsable de reservar memoria para los tres elementos (por ejemplo con `malloc`), `ptr_raiz` apunta a una variable y esta variable apunta a la raíz del árbol. Para generar el primer nodo del árbol (la raíz), se debe asignar `NULL` a la variable; de este modo la variable recibe un puntero a la raíz del árbol.

`tfind` es idéntica a `tsearch`, con la excepción de que si el elemento no se encuentra, se devuelve `NULL` y no se añade ningún nodo al árbol.

`twalk` recorre un árbol o un subárbol, ya que se puede usar cualquier nodo como argumento raíz. Cabe notar que se da un puntero a `raiz`, no un puntero a un puntero como usan `tsearch` y `tfind`. `accion` es la función que se invoca para cada nodo del árbol. Esta función debe ser proporcionada también por el usuario. Recibe tres argumentos: un puntero al elemento del árbol, información acerca de cuándo se visita el elemento, y la profundidad del elemento del árbol (`raiz=0`).

La información de la visita es un tipo enumerado `typedef enum preorder, postorder, endorder, leaf VISIT`; donde `preorder`, `postorder` y `endorder` representan la primera, segunda y tercera visitas durante una búsqueda en profundidad de izquierda a derecha del árbol. `leaf` se refiere a la hoja de un árbol.

Recordatorio

Asegúrate de que los casos de prueba que elabores generan salidas tales que te permitan detectar fallos.

F.29 Código Fuente *ntree* (E00)

```
1  #ifndef NTREEH
2  #define NTREEH

4  struct nodo_arbol {
5      char *clave;
6      char *dato;
7      struct nodo_arbol *padre;
8      int n_hijos;
9      int max_hijos;
10     struct nodo_arbol **hijos;
11 };

13 typedef struct nodo_arbol NODO;

15 struct raiz_arbol {
16     struct nodo_arbol *raiz;
17 };

19 typedef struct raiz_arbol ARBOL;

21 #define CAPACIDAD_INICIAL 4

23 #endif
```

Listado F.5: ntree.h

```
1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <assert.h>
4  #include <string.h>
5  #include "ntree.h"

7  static NODO *inicializar_nodo (char *clave,
8                                char *dato,
9                                NODO *padre)
10 {
11     NODO *nodo;

13     assert(clave != NULL);
14     assert(dato != NULL);
15     nodo = (NODO *)malloc(sizeof(NODO));
16     assert(nodo != NULL);
17     nodo->clave = clave;
18     nodo->dato = dato;
19     nodo->n_hijos = 0;
20     nodo->padre = padre;
21     nodo->max_hijos = CAPACIDAD_INICIAL;
22     nodo->hijos = (NODO **)malloc(sizeof(NODO *) * CAPACIDAD_INICIAL);
23     assert(nodo->hijos != NULL);

25     return nodo;
26 }
```

```

28 ARBOL *crear_arbol(char *clave,
29                   char *dato)
30 {
31     ARBOL *arbol;

33     assert(clave != NULL);
34     assert(dato != NULL);
35     arbol = (ARBOL *)malloc(sizeof(ARBOL));
36     assert(arbol != NULL);
37     arbol->raiz = inicializar_nodo(clave, dato, NULL);
38     assert(arbol->raiz != NULL);

40     return arbol;
41 }

44 static NODO *encontrar_nodo(NODO *nodo,
45                             char *clave)
46 {
47     int i;
48     NODO *encontrado, *resultado;

50     assert(nodo != NULL);
51     assert(clave != NULL);
52     encontrado = NULL;
53     resultado = NULL;
54     if (strcmp(clave, nodo->clave) == 0) {
55         resultado = nodo;
56     }
57     else {
58         for (i = 0; i < nodo->n_hijos; ++i) {
59             encontrado = encontrar_nodo(nodo->hijos[i], clave);
60             if (encontrado)
61                 resultado = encontrado;
62         }
63     }
64     return resultado;
65 }

68 int aniadir_hijo_arbol(ARBOL *arbol,
69                        char *clave_padre,
70                        char *clave_hijo,
71                        char *dato_hijo)
72 {
73     NODO *padre, *hijo;
74     NODO **espacio;
75     int resultado;

77     assert(arbol != NULL);
78     assert(clave_padre != NULL);
79     assert(clave_hijo != NULL);
80     padre = encontrar_nodo(arbol->raiz, clave_padre);
81     if (padre == NULL) {
82         printf("No se encuentra nodo padre con clave %s\n", clave_padre);

```



```
83     resultado = -1;
84 }
85 else {
86     if (padre->n_hijos == padre->max_hijos) {
87         bcopy (padre->hijos, espacio, sizeof(NODO *) * padre->max_hijos);
88         free(padre->hijos);
89         padre->hijos = espacio;
90         (padre->max_hijos) *= 2;
91     }
92     hijo = inicializar_nodo(clave_hijo, dato_hijo, padre);
93     assert(hijo != NULL);
94     padre->hijos[padre->n_hijos] = hijo;
95     ++(padre->n_hijos);
96     resultado = 0;
97 }

99     return resultado;
100 }

103 int buscar_arbol(ARBOL *arbol,
104                 char *clave)
105 {
106     int resultado;
107     NODO *nodo;

109     resultado = -1;
110     assert(arbol != NULL);
111     assert(clave != NULL);
112     nodo = encontrar_nodo(arbol->raiz, clave);
113     if (nodo != NULL) {
114         printf("El contenido es: %s\n", nodo->dato);
115         resultado = 0;
116     } else printf("No se encuentra ningun nodo con clave %s\n", clave);

118     return resultado;
119 }

122 int son_hermanos_arbol(ARBOL *arbol,
123                       char *clave1,
124                       char *clave2)
125 {
126     NODO *nodo1, *nodo2;
127     int resultado;

129     assert(arbol != NULL);
130     assert(clave1 != NULL);
131     assert(clave2 != NULL);
132     resultado = 0;
133     nodo1 = encontrar_nodo(arbol->raiz, clave1);
134     nodo2 = encontrar_nodo(arbol->raiz, clave2);
135     if ((nodo1 == NULL) && (nodo2 == NULL))
136         printf("Los nodos %s y %s no son hermanos\n", clave1, clave2);
137     else if (nodo1 == NULL) {
```

```

138     printf("No se encuentra ningun nodo con clave %s\n", clave1);
139     resultado = -1;
140 } else {
141     printf("Los nodos %s y %s %sson hermanos\n", clave1, clave2,
142           (((nodo1->padre == nodo2->padre) && (nodo1!=nodo2)) ? "" : "no "));
143 }
144 return resultado;
145 }

148 static void imprimir_nodos_arbol(NODO *nodo,
149                                 int nivel)
150 {
151     int i;

153     assert(nodo != (NODO *)0);
154     for (i=0; i<nivel; i++)
155         printf(" ");
156     printf("Nodo (Nivel %d): Clave '%s', Contenido '%s'\n",
157           nivel, nodo->clave, nodo->dato);
158     for (i = 0; i < nodo->n_hijos; ++i) {
159         imprimir_nodos_arbol(nodo->hijos[i], nivel+1);
160     }
161 }

164 int imprimir_arbol(ARBOL *arbol)
165 {
166     int resultado = -1;

168     assert (arbol != NULL);
169     assert (arbol->raiz != NULL);
170     if (arbol->raiz->n_hijos != 0) {
171         imprimir_nodos_arbol(arbol->raiz, 0);
172         resultado = 0;
173     }

175     return resultado;
176 }

179 /*
180  * Esta es la función principal
181  * No hay que crear casos de prueba o construir abstracciones
182  * para nada de lo que hay bajo esta línea
183  */
184 void procesar_comandos(filep)
185     FILE *filep;
186 {
187     char *ptr;
188     char buf[BUFSIZ];
189     char comando[BUFSIZ], arg1[BUFSIZ], arg2[BUFSIZ], arg3[BUFSIZ];
190     ARBOL *miarbol = NULL;

192     while (fgets(buf, sizeof(buf), filep) != NULL) {

```

```
193     if ( (ptr = strchr(buf, '\n')) != NULL)
194         *ptr = '\0';
195     printf("\nProcesando linea '%s':\n", buf);
196     *comando = '\0';
197     *arg1 = '\0';
198     *arg2 = '\0';
199     *arg3 = '\0';
200     sscanf(buf, "%s %s %s %s", comando, arg1, arg2, arg3);
201     if (strcmp(comando, "raiz") == 0)
202         miarbol = crear_arbol(strdup(arg1), strdup(arg2));
203     else if (strcmp(comando, "hijo") == 0)
204         aniadir_hijo_arbol(miarbol, arg1, strdup(arg2), strdup(arg3));
205     else if (strcmp(comando, "buscar") == 0)
206         buscar_arbol(miarbol, arg1);
207     else if (strcmp(comando, "hermanos") == 0)
208         son_hermanos_arbol(miarbol, arg1, arg2);
209     else if (strcmp(comando, "imprimir") == 0)
210         imprimir_arbol(miarbol);
211     else
212         printf("Commando '%s' no reconocido\n", comando);
213 }
214 }

216 int main(argc, argv)
217     int argc; char **argv;
218 {
219     FILE *filep;
220     char *file;

222     if (argc != 2)
223         fprintf(stderr, "Uso: %s Dato\n", *argv);
224     else {
225         file = argv[1];
226         if ((filep = fopen(file, "r")) == NULL)
227             perror(file);
228         else {
229             printf("Procesando fichero de entrada '%s'.\n", file);
230             procesar_comandos(filep);
231             printf("Fin de fichero de entrada '%s'.\n", file);
232             fclose(filep);
233         }
234     }
235     return 0;
236 }
```

Listado F.6: ntree.c

F.30 Especificación del programa ‘ntree’ (E01)

Nombre

`ntree` – Funciones para gestionar un árbol n-ario.

Uso

`ntree` fichero-de-entrada

Descripción

`ntree` lee comandos de un fichero y los procesa para probar una serie de funciones, las cuales permiten manejar el tipo abstracto de datos correspondiente a un árbol n-ario. Es decir un árbol cuyos nodos pueden tener cualquier número de hijos. El número de hijos no está limitado por el programa. Un nodo contiene dos cadenas de caracteres: clave y contenido.

Opciones

No tiene opciones

Datos de Entrada

Sólo se permiten los siguientes comandos en el fichero de entrada. Cada comando debe comenzar en una nueva línea. El uso de mayúsculas y minúsculas se distingue. El separador entre comandos y argumentos es el espacio; lo que significa que no puede haber espacios dentro de una cadena. En esta implementación, las líneas de comando (comando junto sus correspondientes parámetros) leen del fichero y están limitadas a una longitud máxima de 1024.

- **raiz** <clave> <contenido>

Crea e inicializa la raíz del árbol. Este comando debe ser el primero comando del fichero de entrada. Si no aparece como el primer comando, o si el comando aparece repetido, el comportamiento del programa no está definido.

- **hijo** <clave-del-padre> <clave-del-hijo> <contenido-del-hijo>

Asocia el nuevo nodo hijo con los datos proporcionados al nodo padre proporcionado. Si no se encuentra el nodo padre en el árbol, aparecerá el correspondiente mensaje de error. Si la clave del nodo padre aparece más de una vez en el árbol, el nuevo hijo se asocia al primer nodo encontrado (ver también “buscar” en la siguiente línea).

- **buscar** <clave>

Se busca en el árbol de izquierda a derecha usando búsqueda en profundidad, en la cual la clave de un nodo se comprueba en la primera visita (preorden). Si se encuentra un nodo con clave igual a la proporcionada, se imprime el contenido; si no, se genera el correspondiente mensaje de error. Si varios nodos tienen la misma clave, sólo se imprime el contenido del primero que se ha encontrado.

- **hermanos** <clave1> <clave2>

Comprueba si los nodos correspondientes a las claves suministradas son hermanos; es decir, si tienen el mismo nodo padre. Si una de las dos claves no se encuentra en el árbol, se imprime el correspondiente mensaje de error. Cada nodo es hermano de sí mismo. Ver también **buscar** para una descripción del método de búsqueda.

- **imprimir**

La funcionalidad para leer ficheros es para propósito de las pruebas y de este modo no es especialmente tolerante a fallos. Por ejemplo, uno puede asumir que comandos que

están escritos de forma incorrecta, comandos con argumentos que faltan, o comandos con demasiados argumentos no se comprobarán con demasiado detalle.

Limitaciones

La funcionalidad para leer ficheros es para propósito de las pruebas y de este modo no es especialmente tolerante a fallos. Por ejemplo, uno puede asumir que comandos que están escritos de forma incorrecta, comandos con argumentos que faltan, o comandos con demasiados argumentos no se comprobarán con demasiado detalle.

F.31 Suplemento “Revisión de código” y “ntree” (E02)

Ésta es la hoja suplementaria para aplicar la técnica “revisión de código” al programa ntree.

Visión general

El programa lee un fichero que contiene comandos para manejar una estructura de datos correspondiente a un árbol, y ha de procesar dichos comandos de modo apropiado. A continuación aparece un ejemplo sencillo de su uso:

```
C:\ type entrada
raiz Asignatura Ingenieria_del_Software
hijo Asignatura Tema1 Conceptos_basicos
imprimir

C:\ ntree entrada
El fichero de entrada 'entrada' está siendo procesado
La linea 'raiz Asignatura Ingeniería_del_Software' está siendo procesada:
La linea 'hijo Asignatura Tema_1 Conceptos_basicos' está siendo procesada:
La linea 'imprimir' está siendo procesada:
    Nodo (Nivel 0): Clave 'Asignatura', Contenido 'Ingenieria_del_Software'
    Nodo (Nivel 1): Clave 'Tema1', Contenido 'Conceptos_basicos'
Fin de fichero de entrada entrada'
```

Breve descripción de las funciones de librería usadas

- **assert(expresión)**
Macro que espera que la expresión que se le proporciona sea cierta en tiempo de ejecución, y en ese caso no hace nada. En otro caso genera un mensaje de error mostrando la expresión, y el programa termina.
- **char* malloc(unsigned tamaño)**
Devuelve un puntero a un bloque de memoria que es contiene al menos **tamaño** bytes. En caso de fallo, devuelve NULL.
- **int free(ptr)**
Devuelve un bloque de memoria pedido anteriormente al sistema. **ptr** debe haber sido pedido usando **malloc** o una función similar. En caso de éxito, devuelve 1, y si no devuelve 0
- **int strcmp(s1,s2)**
Compara dos cadenas. Devuelve un valor mayor, igual o menor que 0 dependiendo de si **s1** es lexicográficamente hablando mayor, igual o menor que **s2**.
- **void bcopy(b1,b2,longitud)**
Copia **longitud** bytes de la cadena **b1** a la cadena **b2**. Las cadenas que se solapan se tratan correctamente.
- **int sscanf(s,formato,punteros)**
Lee caracteres de la cadena **s**, los convierte al formato indicado y escribe el resultado en las variables a las que apunta **punteros**. Es la función contraria a **printf**; las cadenas de especificación de formato son las mismas.
- **char *fgets(s,n,flujo)**
Lee caracteres de **flujo** y los escribe en la cadena **s**, hasta que se hayan leído **n-1** caracteres o se haya encontrado un salto de línea o un fin de fichero.

- `char *strdup(s1)`

Duplica la cadena `s1`. Reserva memoria para este propósito usando `malloc()` y devuelve un puntero a la cadena duplicada en caso de éxito y si no, devuelve `NULL`.

Recordatorio

No se deben generar abstracciones para la función principal del programa.

F.32 Suplemento “Pruebas funcionales” y “ntree” (E03)

Ésta es la hoja suplementaria para aplicar la técnica “pruebas funcionales” al programa ntree.

Visión general

El programa lee un fichero que contiene comandos para manejar una estructura de datos correspondiente a un árbol, y ha de procesar dichos comandos de modo apropiado. A continuación aparece un ejemplo sencillo de su uso:

```
C:\ type entrada
raiz Asignatura Ingenieria_del_Software
hijo Asignatura Tema1 Conceptos_basicos
imprimir

C:\ ntree entrada
El fichero de entrada 'entrada' está siendo procesado
La línea 'raiz Asignatura Ingeniería_del_Software' está siendo procesada:
La línea 'hijo Asignatura Tema_1 Conceptos_basicos' está siendo procesada:
La línea 'imprimir' está siendo procesada:
    Nodo (Nivel 0): Clave 'Asignatura', Contenido 'Ingenieria_del_Software'
        Nodo (Nivel 1): Clave 'Tema1', Contenido 'Conceptos_basicos'
Fin de fichero de entrada entrada'
```

Documentos y códigos usados

Para este ejercicio es necesario el documento E01, que incluye la especificación del componente, así como el código fuente del programa, que será entregado después de que haber escrito los casos de prueba.

Búsqueda de los ficheros necesarios

Crea un nuevo directorio para evitar conflictos con cualquier otro ejercicio de pruebas y copia el fichero ntree.exe del directorio que se ha indicado.

Elaboración de los casos de prueba

Este programa contiene una función para procesar el fichero de entrada y simplificar el uso del programa, que no es necesario probar. *No crees clases de equivalencia para comandos sintácticamente incorrectos*, ya que eso probaría la función de procesamiento. Ejemplos de casos de prueba no necesarios son:

```
borrar todo (El comando no está definido)
hermanos nodo1 (Falta el segundo parámetro)
```

Recordatorio

Asegúrate de que tus casos de prueba generan salidas tales que te permitan detectar los fallos.

F.33 Suplemento “Pruebas estructurales” y “ntree” (E04)

Ésta es la hoja suplementaria para aplicar la técnica “pruebas estructurales” al programa `ntree`. Este ejercicio prueba varias funciones escritas en C. Para que no malgastes esfuerzos y tiempo en este ejercicio, se ha creado una función llamada `procesar_comandos` que se encarga de leer y procesar el fichero de entrada. No es necesario que pruebes ésta función.

Visión general

El programa lee un fichero que contiene comandos para manejar una estructura de datos correspondiente a un árbol, y ha de procesar dichos comandos de modo apropiado. A continuación aparece un ejemplo sencillo de su uso:

```
C:\ type entrada
raiz Asignatura Ingenieria_del_Software
hijo Asignatura Tema1 Conceptos_basicos
imprimir
C:\ ntree entrada
El fichero de entrada 'entrada' está siendo procesado
La linea 'raiz Asignatura Ingenieria_del_Software' está siendo procesada:
La linea 'hijo Asignatura Tema1 Conceptos_basicos' está siendo procesada:
La linea 'imprimir' está siendo procesada:
    Nodo (Nivel 0): Clave 'Asignatura', Contenido 'Ingenieria_del_Software'
    Nodo (Nivel 1): Clave 'Tema1', Contenido 'Conceptos_basicos'
Fin de fichero de entrada entrada'
```

Documentos y códigos usados

Para este ejercicio es necesario el código fuente del programa así como el documento E01, que contiene la especificación del componente, y que te será entregado después de haber creados los casos de prueba y haber intentado conseguir la cobertura pedida.

Búsqueda de los ficheros necesarios

Crea un nuevo directorio para evitar conflictos con cualquier otro ejercicio de pruebas y copia el fichero `ntree.exe` del directorio indicado.

Creación de los casos de prueba

Todas las funciones de programa se prueban vía la invocación:

```
ntree <fichero_de_entrada>
```

El formato del fichero de entrada es el siguiente:

```
Comando1 parametro1 parametro2
Comando2 parametroX
...
```

Los comandos permitidos en la entrada y las funciones que invocan dichos comandos aparecen en la función `procesar_comandos` en el fichero `ntree.c`. Debido a que esta función se ha creado para facilitar las pruebas y no será parte de la aplicación, no es especialmente robusta. Puedes asumir que la función, aunque es sensible, no tiene faltas. No escribas ningún caso de prueba para ejercitar esta función.

Breve descripción de las funciones de librería usadas

- **assert(expresión)**
Macro que espera que la expresión que se le proporciona sea cierta en tiempo de ejecución, y en ese caso no hace nada. En otro caso genera un mensaje de error mostrando la expresión, y el programa termina.
- **char* malloc(unsigned tamaño)**
Devuelve un puntero a un bloque de memoria que es contiene al menos **tamaño** bytes. En caso de fallo, devuelve NULL.
- **int free(ptr)**
Devuelve un bloque de memoria pedido anteriormente al sistema. **ptr** debe haber sido pedido usando **malloc** o una función similar. En caso de éxito, devuelve 1, y si no devuelve 0
- **int strcmp(s1,s2)**
Compara dos cadenas. Devuelve un valor mayor, igual o menor que 0 dependiendo de si **s1** es lexicográficamente hablando mayor, igual o menor que **s2**.
- **void bcopy(b1,b2,longitud)**
Copia **longitud** bytes de la cadena **b1** a la cadena **b2**. Las cadenas que se solapan se tratan correctamente.
- **int sscanf(s,formato,punteros)**
Lee caracteres de la cadena **s**, los convierte al formato indicado y escribe el resultado en las variables a las que apunta **punteros**. Es la función contraria a **printf**; las cadenas de especificación de formato son las mismas.
- **char *fgets(s,n,flujo)**
Lee caracteres de **flujo** y los escribe en la cadena **s**, hasta que se hayan leído **n-1** caracteres o se haya encontrado un salto de línea o un fin de fichero.
- **char *strdup(s1)**
Duplica la cadena **s1**. Reserva memoria para este propósito usando **malloc()** y devuelve un puntero a la cadena duplicada en caso de éxito y si no, devuelve NULL.

Recordatorio

Asegúrate de que los casos de prueba que elabores generan salidas tales que te permitan detectar fallos.

Apéndice G

Tablas con Información de las Replicaciones usadas en la Evaluación

RESUMEN: En este apéndice se encuentran las tablas con información del conjunto de replicaciones usado para la evaluación. En particular se encuentran las tablas de agrupaciones y de particiones usadas en la evaluación del *generador de particiones* así como la tabla maestra y tablas de particiones usadas en la evaluación del *comparador de replicaciones*.

Id. Exp.	Cons. Efecto	Oper. Efecto	Par constructos de causa	Par de operacionalizaciones de causa
umd82, umd83, umd84	Efectividad	Faltas observadas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
		Cobertura de sent.	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
		Percepción del sujeto	Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
			Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
	Eficiencia	Tasa de detección de faltas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
	Coste	Tiempo de detección de faltas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
	Visibilidad	Faltas observadas/observables	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Sentencias vs. Lec. C. por A.S.
ukl94, ukl95	Efectividad	Faltas observadas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
	Eficiencia	Tasa de detección de faltas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
		Tasa de aislamiento	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
	Coste	Tiempo de detección de faltas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.
	Aislamiento	Faltas aisladas	Func. vs. Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados
			Func. vs. Lec. Código	P.C.E. y A.V.L. vs. Lec. C. por A.S.
			Estruct. vs. Lec. Código	Cob. Predicados vs. Lec. C. por A.S.

Tabla G.1: Tabla de agrupaciones del formulario PCR-3

Id. Exp.	Cons. Efecto	Oper. Efecto	Par constructos de causa	Par de operacionalizaciones de causa
uos97	Efectividad	Faltas observadas	Func. vs. Estruct. Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. y A.V.L. vs. Cob. Decisiones P.C.E. y A.V.L. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.
	Eficiencia	Tasa de detección de faltas	Func. vs. Estruct. Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. y A.V.L. vs. Cob. Decisiones P.C.E. y A.V.L. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.
upm01..11, upv05	Efectividad	Faltas observadas	Func. vs. Estruct.	P.C.E. vs. Cob. Decisiones
upm01..10	Efectividad	Faltas observadas	Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.
upm01..10, upv05, uds05, ort05	Efectividad	Faltas observables	Func. vs. Estruct.	P.C.E. vs. Cob. Decisiones
upm01..09, uds05	Efectividad	Faltas observables	Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.
upm01..10, upv05	Efectividad	Faltas observadas por tipo de falta	Func. vs. Estruct.	P.C.E. vs. Cob. Decisiones
upm01..10	Efectividad	Faltas observadas por tipo de falta	Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.
upm01..09, upv05, uds05, ort05	Efectividad	Faltas observables por tipo de falta	Func. vs. Estruct.	P.C.E. vs. Cob. Decisiones
upm01..09, uds05	Efectividad	Faltas observables por tipo de falta	Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.
upm01-11 upv05, uds05 ort05	Eficiencia	Casos de prueba generados	Func. vs. Estruct. Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. vs. Cob. Decisiones P.C.E. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.
upm01-05 upv05 uds05	Visibilidad	Faltas observadas con CP predefinidos	Func. vs. Estruct. Func. vs. Lec. Código Estruct. vs. Lec. Código	P.C.E. vs. Cob. Decisiones P.C.E. vs. Lec. C. por A.S. Cob. Decisiones vs. Lec. C. por A.S.

Tabla G.2: Tabla de agrupaciones del del formulario PCR-3 (continuación)

Partición	Constructos de efecto	Operacionalizaciones de efecto	Par de constructos de causa	Par de operacionalizaciones de causa	Id. Experimentos	Experimentos por partición
01	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..11, upv05	12
02	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias	umd82..84	3
03	Efectividad	Faltas Observadas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados	ukl94, ukl95, uos97	3
04	Efectividad	Faltas Observadas	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A. S.	upm01..10	10
05	Efectividad	Faltas Observadas	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs. Lec. C. por A.S.	udm82..84, ukl94, ukl95, uos97	6
06	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..10, uos97	11
07	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Sentencias vs. Lec. C. por A.S.	umd82..84	3
08	Efectividad	Faltas Observadas	Estruct. vs Lec. Cód.	Cob. Predicados vs. Lec. C. por A.S.	ukl94, ukl95	2
09	Efectividad	Faltas Observables	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..10, upv05, uds05, ort05	13
10	Efectividad	Faltas Observables	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias	umd82..84	3
11	Efectividad	Faltas Observables	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..09, uds05	10
12	Efectividad	Faltas Observables	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs Lec. C. por A.S.	umd82..84	3
13	Efectividad	Faltas Observables	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..09, uds05	10
14	Efectividad	Faltas Observables	Estruct. vs Lec. Cód.	Cob. Sentencias vs. Lec. C. por A.S.	umd82..84	3
15	Efectividad	Faltas observadas por tipo de falta	Func. vs Estruct.	P.C.E. vs. Cob. Decisiones	upm01..10, upv05	11
16	Efectividad	Faltas observadas por tipo de falta	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..10, upv05	10
17	Efectividad	Faltas observadas por tipo de falta	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..10	10
18	Efectividad	Faltas observables por tipo de falta	Func. vs Estruct.	P.C.E. vs Cob. Decisiones	upm01..09, upv05, uds05, ort05	12
19	Efectividad	Faltas observables por tipo de falta	Func. vs Lec. Cód.	P.C.E. vs Lec. C. por A.S.	upm01..09, uds05	10
20	Efectividad	Faltas observables por tipo de falta	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	upm01..09, uds05	10
21	Efectividad	Cobertura de sentencias	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias	umd82..84	3
22	Efectividad	Cobertura de sentencias	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs. Lec. C. por A.S.	umd82..84	3
23	Efectividad	Cobertura de sentencias	Estruct. vs Lec. Cód.	Cob. Sentencias vs. Lec. C. por A.S.	umd82..84	3

Tabla G.3: Particiones identificadas en el conjunto de replicaciones

Partición	Constructos de efecto	Operacionalizaciones de efecto	Par de constructos de causa	Par de operacionalizaciones de causa	Id. Experimentos	Experimentos por partición
24	Efectividad	Percepción del sujeto	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Sentencias	umd82..84	3
25	Efectividad	Percepción del sujeto	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs. Lec. C. por A.S.	umd82..84	3
26	Efectividad	Percepción del sujeto	Estruct. vs Lec. Cód.	Cob. Sentencias vs. Lec. C. por A.S.	umd82..84	3
27	Eficiencia	Tasa de detección de faltas	Func. vs Estruct.	P.C.E. y A.V.L. vs Cob. Sentencias	umd82..84	3
28	Eficiencia	Tasa de detección de faltas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados	ukl94, ukl95	2
29	Eficiencia	Tasa de detección de faltas	Func. vs Estruct.	P.C.E. y A.V.L. vs Cob. Decisiones	uos97	1
30	Eficiencia	Tasa de detección de faltas	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs Lec. C. por A.S.	umd82..84, ukl94, ukl95, uos97	6
31	Eficiencia	Tasa de detección de faltas	Estruct. vs Lec. Cód.	Cob. Sentencias vs Lec. C. por A.S.	umd82..84	3
32	Eficiencia	Tasa de detección de faltas	Estruct. vs Lec. Cód.	Cob. Predicados vs. Lec. C. por A.S.	ukl94, ukl95	2
33	Eficiencia	Tasa de detección de faltas	Estruct. vs Lec. Cód.	Cob. Decisiones vs Lec. C. por A.S.	uos97	1
34	Eficiencia	Casos de prueba generados	Func. vs Estruct.	P.C.E. vs Cob. Decisiones	upm01..11, upv05, uds05, ort05	14
35	Eficiencia	Tasa de aislamiento de faltas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados	ukl94, ukl95	3
36	Eficiencia	Tasa de aislamiento de faltas	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs Lec. C. por A.S.	ukl94, ukl95	3
37	Eficiencia	Tasa de aislamiento de faltas	Estruct. vs Lec. Cód.	Cob. Predicados vs. Lec. C. por A.S.	ukl94, ukl95	3
38	Coste	Tiempo de detección de faltas	Func. vs Estruct.	P.C.E. y A.V.L. vs Cob. Sentencias	umd82..84	3
39	Coste	Tiempo de detección de faltas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados	ukl94, ukl95	2
40	Coste	Tiempo de detección de faltas	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs Lec. C. por A.S.	umd82..84, ukl94, ukl95	5
41	Coste	Tiempo de detección de faltas	Estruct. vs Lec. Cód.	Cob. Sentencias vs Lec. C. por A.S.	umd82..84	3
42	Coste	Tiempo de detección de faltas	Estruct. vs Lec. Cód.	Cob. Predicados vs. Lec. C. por A.S.	ukl94, ukl95	2
43	Visibilidad	Faltas observadas/ observables	Func. vs Estruct.	P.C.E. y A.V.L. vs Cob. Sentencias	umd82..84	3
44	Visibilidad	Faltas observadas con CP predefinidos	Func. vs Estruct.	P.C.E. vs Cob. Decisiones	upm01..05, upv05, uds05	7
45	Aislamiento	Faltas aisladas	Func. vs Estruct.	P.C.E. y A.V.L. vs. Cob. Predicados	ukl94, ukl95	2
46	Aislamiento	Faltas aisladas	Func. vs Lec. Cód.	P.C.E. y A.V.L. vs Lec. C. por A.S.	ukl94, ukl95	2
47	Aislamiento	Faltas aisladas	Estruct. vs Lec. Cód.	Cob. Predicados vs. Lec. C. por A.S.	ukl94, ukl95	2

Tabla G.4: Particiones identificadas en el conjunto de replicaciones (continuación)

Elemento Principal	Experimentadores y Sitio									
	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04
Diseñador	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Vegas	Vegas	Vegas	Vegas
Entrenador	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Juristo y Vegas	Juristo y Vegas	Juristo y Vegas	Juristo y Vegas
Monitor	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Juristo y Vegas	Juristo, Vegas, Dieste y Ferré	Juristo y Vegas	Juristo y Vegas
Medición	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Vegas	Vegas	Vegas	Vegas
Analista	Basili y Selby	Basili y Selby	Basili y Selby	Kamsties y Lott	Kamsties y Lott	Roper, Wood y Miller	Vegas	Vegas	Vegas	Vegas
Sitio	UMD	UMD	NASA	UKL	UKL	UoS	UPM	UPM	UPM	UPM

Tabla G.5: **Tabla maestra**, elementos identificados en experimentadores y sitio

Experimentadores y Sitio (continuación)										
Elemento Principal	UPV05	UdS05	ORT05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Diseñador	Vegas	Vegas	Solari	Vegas	Vegas y Solari	Vegas y Solari	Vegas y Solari	Gómez	Gómez	Vegas y Apa
Entrenador	Letelier	Ramos	Solari	Juristo y Moreno	Juristo y Moreno	Juristo y Vegas	Juristo y Vegas	Juristo y Vegas	Juristo y Vegas	Juristo y Vegas
Monitor	Letelier	Ramos	Solari	Vegas y Moreno	Vegas y Moreno	Juristo y Vegas	Juristo, Vegas y Gómez	Juristo, Vegas y Gómez	Juristo, Vegas y Gómez	Vegas, Gómez y Apa
Medición	Solari	Solari	Solari	Solari	Solari	Solari	Gómez	Solari	Vegas	Apa
Analista	Vegas y Solari	Vegas y Solari	Vegas y Solari	Vegas	Solari	Solari	Gómez	Gómez	Gómez	Apa
Sitio	UPV	US	ORT	UPM	UPM	UPM	UPM	UPM	UPM	UPM

Tabla G.6: **Tabla maestra**, elementos identificados en experimentadores y sitio (continuación)

Protocolo experimental								
Elemento principal	Sub-elemento	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01
Análisis	Técnica de análisis de datos	ANOVA mixto	ANOVA mixto	ANOVA mixto	ANOVA de una vía	ANOVA de una vía	ANOVA mixto	ANOVA factorial
Diseño	Diseño experimental	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto
	Orden de aplicación de los tratamientos	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden
	Combinación del Entrenamiento con las sesiones	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial
	Tratamientos por sesión	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa
	Sesiones requeridas por sujeto para aplicar tratamiento	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión
	Tareas por sujeto por sesión	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa
Instrumentos	Instrumentos o equipo de medición	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)
Objetos experimentales	Programas	P1, P2, P3	P1, P2, P4	P1, P3, P4	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee
	Faltas (Instancias)	Instancia A	Instancia B	Instancia C	Instancia D	Instancia E	Instancia F	Instancia G y H
Materiales	Formularios del experimento	Formularios v1	Formularios v1	Formularios v1	Formularios v2	Formularios v2	Formularios v2	Formularios v2
	Instrucciones del experimento	Instrucciones v1	Instrucciones v1	Instrucciones v1	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2
	Guías suplementarias	Guías v1	Guías v1	Guías v1	Guías v2	Guías v2	Guías v2	Guías v2

Tabla G.7: **Tabla maestra**, elementos identificados en protocolo experimental

Protocolo experimental (continuación)								
Elemento principal	Sub-elemento	UPM02	UPM03	UPM04	UPV05	UdS05	ORT05	UPM05
Análisis	Técnica de análisis de datos	ANOVA factorial	ANOVA factorial	ANOVA factorial	ANOVA factorial	ANOVA factorial	ANOVA factorial	ANOVA factorial
Diseño	Diseño experimental	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto
	Orden de aplicación de los tratamientos	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	E, F	E, F, R	Todos los tipos de orden	Todos los tipos de orden
	Combinación del Entrenamiento con las sesiones	Secuencial	Secuencial	Secuencial	Intercalado	Intercalado	Secuencial	Secuencial
	Tratamientos por sesión	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	1 Técnica, 3 Programas	1 Técnica, 3 Programas	2 Técnicas, 2 Programas	3 Técnicas, 1 Programa
	Sesiones requeridas por sujeto para aplicar tratamiento	1 Sesión	1 Sesión	1 Sesión	2 Sesiones	2 Sesiones	1 Sesión	1 Sesión
	Tareas por sujeto por sesión	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	2 Técnicas, 2 Programas	1 Técnica, 1 Programa
Instrumentos	Instrumentos o equipo de medición	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)
Objetos experimentales	Programas	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl	Cmdline, Nametbl, Ntee
	Faltas (Instancias)	Instancia G y H	Instancia G y H	Instancia G y H	Instancia G y H	Instancia G y H	Instancia G y H	Instancia G y H
Materiales	Formularios del experimento	Formularios v2	Formularios v2	Formularios v2	Formularios v2	Formularios v2	Formularios v2	Formularios v2
	Instrucciones del experimento	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2
	Guías suplementarias	Guías v2	Guías v2	Guías v2	Guías v2	Guías v2	Guías v2	Guías v2

Tabla G.8: **Tabla maestra**, elementos identificados en protocolo experimental (continuación)

Protocolo experimental (continuación)							
Elemento principal	Sub-elemento	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Análisis	Técnica de análisis de datos	ANOVA factorial	ANOVA factorial	ANOVA factorial	ANOVA factorial	ANOVA factorial	ANOVA factorial
Diseño	Diseño experimental	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial Mixto	Diseño Factorial	Diseño Factorial	Diseño Factorial Mixto
	Orden de aplicación de los tratamientos	Todos los tipos de orden	Todos los tipos de orden	Todos los tipos de orden	No aplica (c/ sujeto aplica una técnica)	No aplica (c/ sujeto aplica una técnica)	E, F
	Combinación del Entrenamiento con las sesiones	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial	Secuencial
	Tratamientos por sesión	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 1 Programa	3 Técnicas, 3 Programas	3 Técnicas, 3 Programas	2 Técnicas, 2 Programas
	Sesiones requeridas por sujeto para aplicar tratamiento	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión	1 Sesión
	Tareas por sujeto por sesión	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa	1 Técnica, 1 Programa
Instrumentos	Instrumentos o equipo de medición	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)	Ordenador, folios, bolígrafos (manual)
Objetos experimentales	Programas	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Cmdline, Nametbl, Ntee	Nametbl, Ntee
	Faltas (Instancias)	Instancia I	Instancia I	Instancia I	Instancia I	Instancia I	Instancia I
Materiales	Formularios del experimento	Formularios v2	Formularios v2	Formularios v2	Formularios v2	Formularios v2	Formularios v2
	Instrucciones del experimento	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2	Instrucciones v2
	Guías suplementarias	Guías v2	Guías v2	Guías v2	Guías v2	Guías v2	Guías v2

Tabla G.9: **Tabla maestra**, elementos identificados en protocolo experimental (continuación)

Operacionalizaciones de los constructos							
Elemento	Sub-elemento	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97
(Causas) Tratamiento A	Tratamiento acorde a la literatura de SE	P.C.E y A.V.L.	P.C.E y A.V.L.	P.C.E y A.V.L.	P.C.E y A.V.L.	P.C.E y A.V.L.	P.C.E y A.V.L.
	Procedimiento de aplicación del tratamiento	Procedimiento de aplicación P.C.E. v1	Procedimiento de aplicación P.C.E. v1	Procedimiento de aplicación P.C.E. v1	Procedimiento de aplicación P.C.E. v2	Procedimiento de aplicación P.C.E. v2	Procedimiento de aplicación P.C.E. v2
	Tratamiento como instrucciones	P.C.E. v1	P.C.E. v1	P.C.E. v1	P.C.E. v2	P.C.E. v2	P.C.E. v2
	Recursos del tratamiento	Ordenador (1 sujeto por ordenador), Test Driver, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Test Driver, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Test Driver, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Impresora, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Impresora, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Herramienta de cobertura de código
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en tutorial	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario
Tratamiento B	Tratamiento acorde a la literatura de SE	Cobertura de sentencias	Cobertura de sentencias	Cobertura de sentencias	Cobertura de predicados	Cobertura de predicados	Cobertura de decisiones
	Procedimiento de aplicación del tratamiento	Procedimiento cobertura de sentencias	Procedimiento cobertura de sentencias	Procedimiento cobertura de sentencias	Procedimiento cobertura de predicados	Procedimiento cobertura de predicados	Procedimiento cobertura de predicados
	Tratamiento como instrucciones	Cobertura de sentencias v1	Cobertura de sentencias v1	Cobertura de sentencias v1	Cobertura de predicados v1	Cobertura de predicados v1	Cobertura de decisiones v1
	Recursos del tratamiento	Ordenador (1 sujeto por ordenador), Test Driver, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Test Driver, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Test Driver, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Impresora, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Impresora, Herramienta de cobertura de código	Ordenador (1 sujeto por ordenador), Herramienta de cobertura de código
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en tutorial	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario
Tratamiento C	Tratamiento acorde a la literatura de SE	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas
	Procedimiento de aplicación del tratamiento	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1
	Tratamiento como instrucciones	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1
	Recursos del tratamiento	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en tutorial	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario
(Efecto) Efectividad	Métrica para faltas observadas	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Número de faltas observadas por sujeto
		Número de faltas observadas por sujeto	Número de faltas observadas por sujeto	Número de faltas observadas por sujeto			
	Métrica para faltas observables	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto			
	Procedimiento de medición (criterio de corrección)	Criterio relajado	Criterio relajado	Criterio relajado	Criterio relajado	Criterio relajado	Criterio relajado

Tabla G.10: **Tabla maestra**, elementos identificados en operacionalizaciones de los constructos

Operacionalizaciones de los constructos (continuación)								
Elemento	Sub-elemento	UPM01	UPM02	UPM03	UPM04	UPV05	UdS05	ORT05
(Causas) Tratamiento A	Tratamiento acorde a la literatura de SE	P.C.E.	P.C.E.	P.C.E.	P.C.E.	P.C.E.	P.C.E.	P.C.E.
	Procedimiento de aplicación del tratamiento	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3
	Tratamiento como instrucciones	P.C.E. v3	P.C.E. v3	P.C.E. v3	P.C.E. v3	P.C.E. v3	P.C.E. v3	P.C.E. v3
	Recursos del tratamiento	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario y Sin experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en tutorial	Sin experiencia en tutorial	Con experiencia en curso universitario
Tratamiento B	Tratamiento acorde a la literatura de SE	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones
	Procedimiento de aplicación del tratamiento	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones
	Tratamiento como instrucciones	Cobertura de decisiones v2	Cobertura de decisiones v2	Cobertura de decisiones v2	Cobertura de decisiones v2	Cobertura de decisiones v2	Cobertura de decisiones v2	Cobertura de decisiones v2
	Recursos del tratamiento	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario y Sin experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en tutorial	Sin experiencia en tutorial	Con experiencia en curso universitario
Tratamiento C	Tratamiento acorde a la literatura de SE	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas		Lectura de código por abstracciones sucesivas	
	Procedimiento de aplicación del tratamiento	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1		Procedimiento lectura código v1	
	Tratamiento como instrucciones	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1		lectura de código v1	
	Recursos del tratamiento	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos		Folios, bolígrafos	
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario y Sin experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario			
(Efecto) Efectividad	Métrica para faltas observadas	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto		
	Métrica para faltas observables	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto
	Procedimiento de medición (criterio de corrección)	Criterio relajado	Criterio relajado	Criterio relajado	Criterio relajado	Criterio relajado	Criterio relajado	Criterio relajado

Tabla G.11: **Tabla maestra**, elementos identificados en operacionalizaciones de los constructos (continuación)

Operacionalizaciones de los constructos (continuación)								
Elemento	Sub-elemento	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
(Causas) Tratamiento A	Tratamiento acorde a la literatura de SE	P.C.E.	P.C.E.	P.C.E.	P.C.E.	P.C.E.	P.C.E.	P.C.E.
	Procedimiento de aplicación del tratamiento	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3	Procedimiento de aplicación P.C.E. v3
	Tratamiento como instrucciones	P.C.E. v3	P.C.E. v4	P.C.E. v4	P.C.E. v4	P.C.E. v4	P.C.E. v4	P.C.E. v4
	Recursos del tratamiento	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario
Tratamiento B	Tratamiento acorde a la literatura de SE	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones	Cobertura de decisiones
	Procedimiento de aplicación del tratamiento	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones	Procedimiento cobertura de decisiones
	Tratamiento como instrucciones	Cobertura de decisiones v2	Cobertura de decisiones v3	Cobertura de decisiones v3	Cobertura de decisiones v3	Cobertura de decisiones v3	Cobertura de decisiones v3	Cobertura de decisiones v3
	Recursos del tratamiento	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)	Ordenador (1 sujeto por ordenador)
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario
Tratamiento C	Tratamiento acorde a la literatura de SE	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas	Lectura de código por abstracciones sucesivas
	Procedimiento de aplicación del tratamiento	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1	Procedimiento lectura código v1
	Tratamiento como instrucciones	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1	lectura de código v1
	Recursos del tratamiento	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos	Folios, bolígrafos
	Aspectos de transmisión del tratamiento	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario	Con experiencia en curso universitario
(Efecto) Efectividad	Métrica para faltas observadas	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto	Porcentaje de faltas observadas por sujeto
	Métrica para faltas observables	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto	Porcentaje de faltas observables por sujeto
	Procedimiento de medición (criterio de corrección)	Criterio relajado	Criterio relajado	Criterio relajado	Criterio estricto	Criterio relajado	Criterio relajado	Criterio relajado

Tabla G.12: **Tabla maestra**, elementos identificados en operacionalizaciones de los constructos (continuación)

		Poblaciones						
Elemento	Sub-elemento	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01
Propiedades de los sujetos	Tipo de sujeto	Pregrado avanzado, Grado (Máster o Doctorado)	Pregrado avanzado, Grado (Máster o Doctorado)	Profesional	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado
	Experiencia en técnica (tratamiento)	Solo para experimento	Solo para experimento	La usa habitualmente (familiarizado)	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento
	Otras condiciones de los sujetos	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimiento limitado de programación	Conocimiento limitado de programación	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento
Propiedades de los objetos experimentales	Lenguaje	Lenguaje Simplt	Lenguaje Simplt	Lenguaje Fortran	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C
	Tamaño	Más pequeño (< 150 LOC), Pequeño (150-220 LOC)	Más pequeño (< 150 LOC), Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Más pequeño (< 150 LOC), Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)
	Complejidad ciclomática	Baja (< 15), Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Baja (< 15), Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)
	Tipos de Faltas	Grupo de Faltas A	Grupo de Faltas B	Grupo de Faltas C	Grupo de Faltas D	Grupo de Faltas E	Grupo de Faltas F	Grupo de Faltas G, Grupo de Faltas H

Tabla G.13: **Tabla maestra**, elementos identificados en poblaciones

Poblaciones (continuación)								
Elemento	Sub-elemento	UPM02	UPM03	UPM04	UPV05	UdS05	ORT05	UPM05
Propiedades de los sujetos	Tipo de sujeto	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado
	Experiencia en técnica (tratamiento)	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento
	Otras condiciones de los sujetos	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimiento limitado de programación	Conocimientos de prog aptos para experimento
Propiedades de los objetos experimentales	Lenguaje	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C
	Tamaño	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)
	Complejidad ciclomática	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)
	Tipos de Faltas	Grupo de Faltas G, Grupo de Faltas H	Grupo de Faltas G, Grupo de Faltas H	Grupo de Faltas G, Grupo de Faltas H	Grupo de Faltas G, Grupo de Faltas H	Grupo de Faltas G, Grupo de Faltas H	Grupo de Faltas G, Grupo de Faltas H	Grupo de Faltas G, Grupo de Faltas H

Tabla G.14: **Tabla maestra**, elementos identificados en poblaciones (continuación)

Poblaciones (continuación)							
Elemento	Sub-elemento	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Propiedades de los sujetos	Tipo de sujeto	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado	Pregrado avanzado, Grado (Máster o Doctorado)	Pregrado avanzado, Grado (Máster o Doctorado)
	Experiencia en técnica (tratamiento)	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento	Solo para experimento
	Otras condiciones de los sujetos	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento	Conocimientos de prog aptos para experimento
Propiedades de los objetos experimentales	Lenguaje	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C	Lenguaje C
	Tamaño	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)	Pequeño (150-220 LOC), Menos pequeño (221-300 LOC)
	Complejidad ciclométrica	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)	Media baja (15-29), Media (30-45)
	Tipos de Faltas	Grupo de Faltas I	Grupo de Faltas I	Grupo de Faltas I	Grupo de Faltas I	Grupo de Faltas I	Grupo de Faltas I

Tabla G.15: **Tabla maestra**, elementos identificados en poblaciones (continuación)

Sección	Elemento principal	Sub elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05
Experimentadores	Diseñador		1.1	Vegas	X	X	X	X	X
			1.2	Juristo	X	X	X	X	X
	Entrenador		1.2	Vegas	X	X	X	X	
			1.2	Moreno					X
			1.3	Juristo	X	X	X	X	
	Monitor		1.3	Vegas	X	X	X	X	X
			1.3	Dieste		X			
			1.3	Ferre		X			
			1.3	Moreno					X
	Medición		1.4	Vegas	X	X	X	X	
			1.4	Solari					X
	Analista		1.5	Vegas	X	X	X	X	X
Sitio			2	UPM	X	X	X	X	X
Protocolo experimental	Análisis	Técnica de análisis de datos	3.1.1	ANOVA factorial	X	X	X	X	X
	Diseño	Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X	X	X
		Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X	X	X
		Combinación del entrenamiento con las sesiones	3.2.3	Secuencial	X	X	X	X	X
		Tratamientos por sesión	3.2.4	3 Técnicas, 1 Programa	X	X	X	X	X
		Sesiones requeridas por sujeto para aplicar tratamientos	3.2.5	1 Sesión	X	X	X	X	X
		Tareas por sujeto por sesión	3.2.6	1 Técnica, 1 Programa	X	X	X	X	X
	Instrumentos	Instrumentos o equipo de medición	3.3.1	Ordenador, folios, bolígrafos (manual)	X	X	X	X	X
	Objetos experimentales	Programas	3.4.1	Cmdline	X	X	X	X	X
			3.4.1	Nametbl	X	X	X	X	X
			3.4.1	Ntree	X	X	X	X	X
		Faltas (Instancias)	3.4.2	Instancia G	X	X	X	X	X
			3.4.2	Instancia H	X	X	X	X	X
	Materiales	Formularios del experimento	3.5.1	Formularios v2	X	X	X	X	X
		Instrucciones del experimento	3.5.2	Instrucciones v2	X	X	X	X	X
		Guías suplementarias	3.5.3	Guías v2	X	X	X	X	X

Tabla G.16: **Partición 01.a**, tabla referente a experimentadores, sitio y protocolo experimental

					Operacionalizaciones de los constructos					
Elemento principal	Sub-elemento	ID	Categoría		UPM01	UPM02	UPM03	UPM04	UPM05	
(Causas)	Tratamiento A	Tratamiento acorde a la literatura de SE	4.1.1.1	P.C.E.	X	X	X	X	X	
		Procedimiento de aplicación del tratamiento	4.1.1.2	Procedimiento de aplicación P.C.E. v3	X	X	X	X	X	
		Tratamiento como instrucciones	4.1.1.3	P.C.E. v3	X	X	X	X	X	
		Recursos del tratamiento	4.1.1.4	Ordenador (1 sujeto x ordenador)	X	X	X	X	X	
		Aspectos de transmisión del tratamiento	4.1.1.5	Con experiencia en curso universitario	X	X	X	X	X	
	Tratamiento B		4.1.1.5	Sin experiencia en curso universitario	X					
		Tratamiento acorde a la literatura de SE	4.1.2.1	Cobertura de decisiones	X	X	X	X	X	X
		Prodedimiento de aplicación del tratamiento	4.1.2.2	Procedimiento cobertura de decisiones	X	X	X	X	X	X
		Tratamiento como instrucciones	4.1.2.3	Cobertura de decisiones v2	X	X	X	X	X	X
		Recursos del tratamiento	4.1.2.4	Ordenador (1 persona x ord)	X	X	X	X	X	X
		Aspectos de transmisión del tratamiento	4.1.2.5	Con experiencia en curso universitario	X	X	X	X	X	X
			4.1.2.5	Sin experiencia en curso universitario	X					
	(Efectos)	Métrica	4.2.1	Porcentaje de faltas observadas por sujeto	X	X	X	X	X	
	Efectividad	Procedimiento de medición (criterio de corrección)	4.2.2	Criterio relajado	X	X	X	X	X	
Poblaciones										
Elemento principal	Sub-elemento	ID	Categoría		UPM01	UPM02	UPM03	UPM04	UPM05	
Propiedades de los sujetos	Tipo de sujeto	5.1.1	Pregado avanzado		X	X	X	X	X	
	Experiencia en técnica (tratamiento)	5.1.2	Solo para experimento		X	X	X	X	X	
	Otras condiciones de los sujetos	5.1.3	Conocimientos de prog aptos p/ exp.		X	X	X	X	X	
Propiedades de los objetos experimentales	Lenguaje	5.2.1	Lenguaje C		X	X	X	X	X	
	Tamaño	5.2.2	Pequeño (150-220 LOC)		X	X	X	X	X	
		5.2.2	Menos pequeño (221-300 LOC)		X	X	X	X	X	
	Complejidad ciclomática	5.2.3	Media baja 15-29		X	X	X	X	X	
		5.2.3	Alta >45		X	X	X	X	X	
	Tipos de Faltas	5.2.4	Grupo de Faltas G		X	X	X	X	X	
		5.2.4	Grupo de Faltas H		X	X	X	X	X	

Tabla G.17: **Partición 01.a**, tabla referente a operacionalizaciones y poblaciones

Estructura	Elemento principal	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10
Experimentadores	Diseñador	1.1	Vegas	X	X	X	X	X	X	X	X		
		1.1	Gómez									X	X
		1.1	Solari						X	X	X		
	Entrenador	1.2	Juristo	X	X	X	X	X	X	X	X	X	X
		1.2	Vegas	X	X	X	X			X	X	X	X
		1.2	Moreno					X	X				
	Monitor	1.3	Juristo	X	X	X	X			X	X	X	X
		1.3	Vegas	X	X	X	X	X	X	X	X	X	X
		1.3	Dieste		X								
		1.3	Ferre		X								
		1.3	Moreno					X	X				
		1.3	Gómez								X	X	X
	Medición	1.4	Vegas	X	X	X	X						X
		1.4	Gómez								X		
		1.4	Solari					X	X	X		X	
	Analista	1.5	Vegas	X	X	X	X	X					
		1.5	Gómez								X	X	X
		1.5	Solari						X	X			
Sitio		2	UPM	X	X	X	X	X	X	X	X	X	X

Tabla G.18: **Partición 06.b**, tabla referente a los experimentadores y sitio

				Protocolo experimental									
Elemento principal	Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10
Análisis	Técnica de análisis de datos	3.1.1	ANOVA factorial	X	X	X	X	X	X	X	X	X	X
Diseño	Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X	X	X	X	X	X		
		3.2.1	Diseño Factorial									X	X
	Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X	X	X	X	X	X		
		3.2.2	No Aplica (c/ suj app 1 tec)									X	X
	Combinación del entrenamiento con las sesiones	3.2.3	Secuencial	X	X	X	X	X	X	X	X	X	X
	Tratamientos por sesión	3.2.4	3 Técnicas, 1 Programa	X	X	X	X	X	X	X	X		
		3.2.4	3 Técnicas, 3 Programas									X	X
	Sesiones requeridas por sujeto para aplicar tratamientos	3.2.5	1 Sesión	X	X	X	X	X	X	X	X	X	X
Instrumentos	Tareas por sujeto por sesión	3.2.6	1 Técnica, 1 Programa	X	X	X	X	X	X	X	X	X	X
	Instrumentos o equipo de medición	3.3.1	Ordenador, folios, bolígrafos (manual)	X	X	X	X	X	X	X	X	X	X
Objetos experimentales	Programas	3.4.1	Cmdline	X	X	X	X	X	X	X	X	X	X
		3.4.1	Nametbl	X	X	X	X	X	X	X	X	X	X
		3.4.1	Ntree	X	X	X	X	X	X	X	X	X	X
	Faltas (Instancias)	3.4.2	Instancia G	X	X	X	X	X					
		3.4.2	Instancia H	X	X	X	X	X					
		3.4.2	Instancia I						X	X	X	X	X
Materiales	Formularios del experimento	3.5.1	Formularios v2	X	X	X	X	X	X	X	X	X	X
	Instrucciones del experimento	3.5.2	Instrucciones v2	X	X	X	X	X	X	X	X	X	X
	Guías suplementarias	3.5.3	Guías v2	X	X	X	X	X	X	X	X	X	X

Tabla G.19: **Partición 06.b**, tabla referente al protocolo experimental

Operacionalizaciones de los constructos														
Elemento principal	Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	
(Causas) Tratamiento B	Tratamiento acorde a la literatura de SE	4.1.1.1	Cobertura de decisiones	X	X	X	X	X	X	X	X	X	X	
	Procedimiento de aplicación del tratamiento	4.1.1.2	Procedimiento cobertura de decisiones	X	X	X	X	X	X	X	X	X	X	
	Tratamiento como instrucciones	4.1.1.3	Cobertura de decisiones v2	X	X	X	X	X						
		4.1.1.3	Cobertura de decisiones v3						X	X	X	X	X	
	Recursos del tratamiento	4.1.1.4	Ordenador (1 persona x ord)	X	X	X	X	X	X	X	X	X	X	
	Aspectos de transmisión del tratamiento	4.1.1.5	Con experiencia en curso universitario	X	X	X	X	X	X	X	X	X	X	
		4.1.1.5	Sin experiencia en curso universitario	X										
Tratamiento C	Tratamiento acorde a la literatura de SE	4.1.2.1	Lec. de código por abstracciones sucesivas	X	X	X	X	X	X	X	X	X	X	
	Procedimiento de aplicación del tratamiento	4.1.2.2	Procedimiento lectura código v1	X	X	X	X	X	X	X	X	X	X	
	Tratamiento como instrucciones	4.1.2.3	Lectura de código v1	X	X	X	X	X	X	X	X	X	X	
	Recursos del tratamiento	4.1.2.4	Folios, bolígrafos	X	X	X	X	X	X	X	X	X	X	
	Aspectos de transmisión del tratamiento	4.1.2.5	Con experiencia en curso universitario	X	X	X	X	X	X	X	X	X	X	
		4.1.2.5	Sin experiencia en curso universitario	X										
	(Efectos)	Métrica	4.2.1	Porcentaje de faltas observadas por sujeto	X	X	X	X	X	X	X	X	X	X
Efectividad	Procedimiento de medición (criterio de corrección)	4.2.2	Criterio relajado	X	X	X	X	X	X	X		X	X	
		4.2.2	Criterio estricto									X		
Poblaciones														
Elemento principal	Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	
Propiedades de los sujetos	Tipo de sujeto	5.1.1	Pregrado avanzado	X	X	X	X	X	X	X	X	X	X	
		5.1.1	Grado (Máster o Doc.)										X	
	Experiencia en técnica (tratar	5.1.2	Solo para experimento	X	X	X	X	X	X	X	X	X	X	
		5.1.3	Conocimientos de prog aptos p/ exp.	X	X	X	X	X	X	X	X	X	X	
Propiedades de los objetos experimentales	Lenguaje	5.2.1	Lenguaje C	X	X	X	X	X	X	X	X	X	X	
	Tamaño	5.2.2	Pequeño (150-220 LOC)	X	X	X	X	X	X	X	X	X	X	
		5.2.2	Menos pequeño (221-300 LOC)	X	X	X	X	X	X	X	X	X	X	
	Complejidad ciclométrica	5.2.3	Media baja 15-29	X	X	X	X	X	X	X	X	X	X	
		5.2.3	Alta >45	X	X	X	X	X	X	X	X	X	X	
	Tipos de Faltas	5.2.4	Grupo de Faltas G	X	X	X	X	X						
		5.2.4	Grupo de Faltas H	X	X	X	X	X						
		5.2.4	Grupo de Faltas I						X	X	X	X	X	

Tabla G.20: **Partición 06.b**, tabla referente a las operacionalizaciones y poblaciones

Estructura	Elemento principal	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09
Experimentadores	Diseñador	1.1	Vegas	X	X	X	X	X	X	X	X	
		1.1	Gómez									X
		1.1	Solari						X	X	X	
	Entrenador	1.2	Juristo	X	X	X	X	X	X	X	X	X
		1.2	Vegas	X	X	X	X			X	X	X
		1.2	Moreno					X	X			
	Monitor	1.3	Juristo	X	X	X	X			X	X	X
		1.3	Vegas	X	X	X	X	X	X	X	X	X
		1.3	Dieste		X							
		1.3	Ferre		X							
		1.3	Moreno					X	X			
		1.3	Gómez								X	X
	Medición	1.4	Vegas	X	X	X	X					
		1.4	Gómez								X	
		1.4	Solari					X	X	X		X
	Analista	1.5	Vegas	X	X	X	X	X				
		1.5	Gómez								X	X
		1.5	Solari						X	X		
Sitio		2	UPM	X	X	X	X	X	X	X	X	X

Tabla G.21: **Partición 11.c**, tabla referente a los experimentadores y sitio

				Protocolo experimental								
Elemento principal	Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09
Análisis	Técnica de análisis de datos	3.1.1	ANOVA factorial	X	X	X	X	X	X	X	X	X
Diseño	Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X	X	X	X	X	X	
		3.2.1	Diseño Factorial									X
	Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X	X	X	X	X	X	
		3.2.2	No Aplica (c/ suj app 1 tec)									X
	Combinación del entrenamiento con las sesiones	3.2.3	Secuencial	X	X	X	X	X	X	X	X	X
	Tratamientos por sesión	3.2.4	3 Técnicas, 1 Programa	X	X	X	X	X	X	X	X	
		3.2.4	3 Técnicas, 3 Programas									X
	Sesiones requeridas por sujeto para aplicar tratamientos	3.2.5	1 Sesión	X	X	X	X	X	X	X	X	X
	Tareas por sujeto por sesión	3.2.6	1 Técnica, 1 Programa	X	X	X	X	X	X	X	X	X
Instrumentos	Instrumentos o equipo de medición	3.3.1	Ordenador, folios, bolígrafos (manual)	X	X	X	X	X	X	X	X	X
Objetos experimentales	Programas	3.4.1	Cmdline	X	X	X	X	X	X	X	X	X
		3.4.1	Nametbl	X	X	X	X	X	X	X	X	X
		3.4.1	Ntree	X	X	X	X	X	X	X	X	X
	Faltas (Instancias)	3.4.2	Instancia G	X	X	X	X	X				
		3.4.2	Instancia H	X	X	X	X	X				
		3.4.2	Instancia I						X	X	X	X
Materiales	Formularios del experimento	3.5.1	Formularios v2	X	X	X	X	X	X	X	X	X
	Instrucciones del experimento	3.5.2	Instrucciones v2	X	X	X	X	X	X	X	X	X
	Guías suplementarias	3.5.3	Guías v2	X	X	X	X	X	X	X	X	X

Tabla G.22: **Partición 11.c**, tabla referente al protocolo experimental

Operacionalizaciones de los constructos													
Elemento principal		Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09
(Causas)	Tratamiento A	Tratamiento acorde a la literatura de SE	4.1.1.1	P.C.E.	X	X	X	X	X	X	X	X	X
		Procedimiento de aplicación del tratamiento	4.1.1.2	Procedimiento de aplicación P.C.E. v3	X	X	X	X	X	X	X	X	X
		Tratamiento como instrucciones	4.1.1.3	P.C.E. v3	X	X	X	X	X	X	X	X	X
		Recursos del tratamiento	4.1.1.4	Ordenador (1 sujeto x ordenador)	X	X	X	X	X	X	X	X	X
		Aspectos de transmisión del tratamiento	4.1.1.5	Con experiencia en curso universitario	X	X	X	X	X	X	X	X	X
			4.1.1.5	Sin experiencia en curso universitario	X								
	Tratamiento C	Tratamiento acorde a la literatura de SE	4.1.2.1	Lec. de código por abstracciones sucesivas	X	X	X	X	X	X	X	X	X
		Prodedimiento de aplicación del tratamiento	4.1.2.2	Procedimiento lectura código v1	X	X	X	X	X	X	X	X	X
		Tratamiento como instrucciones	4.1.2.3	Lectura de código v1	X	X	X	X	X	X	X	X	X
		Recursos del tratamiento	4.1.2.4	Folios, bolígrafos	X	X	X	X	X	X	X	X	X
		Aspectos de transmisión del tratamiento	4.1.2.5	Con experiencia en curso universitario	X	X	X	X	X	X	X	X	X
			4.1.2.5	Sin experiencia en curso universitario	X								
(Efectos)	Métrica	4.2.1	Porcentaje de faltas observables por sujeto	X	X	X	X	X	X	X	X	X	
Efectividad	Procedimiento de medición (criterio de corrección)	4.2.2	Criterio relajado	X	X	X	X	X	X	X		X	
		4.2.2	Criterio estricto									X	
Poblaciones													
Elemento principal		Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPM05	UPM06	UPM07	UPM08	UPM09
Propiedades de los sujetos	Tipo de sujeto	5.1.1	Pregrado avanzado		X	X	X	X	X	X	X	X	X
	Experiencia en técnica (tratamiento)	5.1.2	Solo para experimento		X	X	X	X	X	X	X	X	X
	Otras condiciones de los sujetos	5.1.3	Conocimientos de prog aptos p/ exp.		X	X	X	X	X	X	X	X	X
Popiedades de los objetos experimentales	Lenguaje	5.2.1	Lenguaje C		X	X	X	X	X	X	X	X	X
	Tamaño	5.2.2	Pequeño (150-220 LOC)		X	X	X	X	X	X	X	X	X
		5.2.2	Menos pequeño (221-300 LOC)		X	X	X	X	X	X	X	X	X
	Complejidad ciclomática	5.2.3	Media baja 15-29		X	X	X	X	X	X	X	X	X
		5.2.3	Alta >45		X	X	X	X	X	X	X	X	X
	Tipos de Faltas	5.2.4	Grupo de Faltas G		X	X	X	X	X				
		5.2.4	Grupo de Faltas H		X	X	X	X	X				
		5.2.4	Grupo de Faltas I							X	X	X	X

Tabla G.23: **Partición 11.c**, tabla referente a operacionalizaciones y poblaciones

Estructura	Elemento principal	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPV05	UdS05	ORT05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10
Experimentadores	Diseñador	1.1	Vegas	X	X	X	X	X	X		X	X	X	X		
		1.1	Gómez												X	X
		1.1	Solari							X		X	X	X		
	Entrenador	1.2	Juristo	X	X	X	X				X	X	X	X	X	X
		1.2	Vegas	X	X	X	X						X	X	X	X
		1.2	Letelier					X								
		1.2	Ramos						X							
		1.2	Moreno								X	X				
		1.2	Solari							X						
	Monitor	1.3	Juristo	X	X	X	X						X	X	X	X
		1.3	Vegas	X	X	X	X				X	X	X	X	X	X
		1.3	Dieste		X											
		1.3	Ferre		X											
		1.3	Moreno								X	X				
		1.3	Gómez											X	X	X
		1.3	Letelier					X								
		1.3	Ramos						X							
		1.3	Solari							X						
	Medición	1.4	Vegas	X	X	X	X									X
		1.4	Gómez											X		
		1.4	Solari					X	X	X	X	X	X		X	
	Analista	1.5	Vegas	X	X	X	X	X	X	X	X					
		1.5	Gómez											X	X	X
		1.5	Solari					X	X	X		X	X			
Sitio		2	UPM	X	X	X	X				X	X	X	X	X	X
		2	UPV					X								
		2	UdS						X							
		2	ORT							X						

Tabla G.24: **Partición 09**, tabla referente a los experimentadores y sitio

													Protocolo experimental				
Elemento principal	Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPV05	US05	ORT05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	
Análisis	Técnica de análisis de datos	3.1.1	ANOVA factorial	X	X	X	X	X	X	X	X	X	X	X	X	X	
Diseño	Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X	X	X	X	X	X	X	X	X			
		3.2.1	Diseño Factorial													X	X
	Orden de aplicación de los tratamientos	3.2.2	Todos los tipos de orden	X	X	X	X				X	X	X	X	X		
		3.2.2	E , F					X									
		3.2.2	E, F, R							X							
		3.2.2	No Aplica (c/ suj app 1 tec)													X	X
	Combinación del entrenamiento con las sesiones	3.2.3	Intercalado					X	X								
		3.2.3	Secuencial	X	X	X	X				X	X	X	X	X	X	X
	Tratamientos por sesión	3.2.4	3 Técnicas, 1 Programa	X	X	X	X					X	X	X	X		
		3.2.4	1 Técnica, 3 Programas					X	X								
		3.2.4	3 Técnicas, 3 Programas													X	X
		3.2.4	2 Técnicas, 2 Progrmas								X						
	Sesiones requeridas por sujeto para aplicar tratamientos	3.2.5	1 Sesión	X	X	X	X				X	X	X	X	X	X	X
		3.2.5	2 Sesiones					X	X								
	Tareas por sujetopor sesión	3.2.6	1 Técnica, 1 Programa	X	X	X	X	X	X	X		X	X	X	X	X	X
		3.2.6	2 Técnicas, 2 Programas								X						
Instrumentos	Instrumentos o equipo de medición	3.3.1	Ordenador, folios, bolígrafos (manual)	X	X	X	X	X	X	X	X	X	X	X	X	X	
Objetos experimentales	Programas	3.4.1	Cmdline	X	X	X	X	X	X	X	X	X	X	X	X	X	
		3.4.1	Nametbl	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		3.4.1	Ntree	X	X	X	X	X	X		X	X	X	X	X	X	X
	Faltas (Instancias)	3.4.2	Instancia G	X	X	X	X	X	X	X	X						
		3.4.2	Instancia H	X	X	X	X	X	X	X	X						
		3.4.2	Instancia I										X	X	X	X	X
Materiales	Formularios del experimento	3.5.1	Formularios v2	X	X	X	X	X	X	X	X	X	X	X	X	X	
	Instrucciones del experimento	3.5.2	Instrucciones v2	X	X	X	X	X	X	X	X	X	X	X	X	X	
	Guías suplementarias	3.5.3	Guías v2	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Tabla G.25: **Partición 09**, tabla referente al protocolo experimental

Operacionalizaciones de los constructos																		
Elemento principal		Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPV05	US05	ORT05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	
(Causas)	Tratamiento A	Tratamiento acorde a la literatura de SE	4.1.1.1	P.C.E.	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Procedimiento de aplicación del tratamiento	4.1.1.2	Procedimiento de aplicación P.C.E. v3	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Tratamiento como instrucciones	4.1.1.3	P.C.E. v3	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Recursos del tratamiento	4.1.1.4	Ordenador (1 sujeto x ordenador)	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Aspectos de transmisión del tratamiento	4.1.1.5	Con experiencia en curso universitario	X	X	X	X			X	X	X	X	X	X	X	
			4.1.1.5	Sin experiencia en curso universitario	X													
			4.1.1.5	Con experiencia en tutorial					X									
			4.1.1.5	Sin experiencia en tutorial						X								
	Tratamiento B	Tratamiento acorde a la literatura de SE	4.1.2.1	Cobertura de decisiones	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Prodedimiento de aplicación del tratamiento	4.1.2.2	Procedimiento cobertura de decisiones	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Tratamiento como instrucciones	4.1.2.3	Cobertura de decisiones v2	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Recursos del tratamiento	4.1.2.4	Ordenador (1 persona x ord)	X	X	X	X	X	X	X	X	X	X	X	X	X	
		Aspectos de transmisión del tratamiento	4.1.2.5	Con experiencia en curso universitario	X	X	X	X			X	X	X	X	X	X	X	
			4.1.2.5	Sin experiencia en curso universitario	X													
			4.1.2.5	Con experiencia en tutorial					X									
			4.1.2.5	Sin experiencia en tutorial						X								
(Efectos)	Métrica	4.2.1	Porcentaje de faltas observables por sujeto	X	X	X	X	X	X	X	X	X	X	X	X	X		
Efectividad	Procedimiento de medición (criterio de corrección)	4.2.2	Criterio relajado	X	X	X	X	X	X	X	X	X	X		X	X		
		4.2.2	Criterio estricto												X			

Tabla G.26: **Partición 09**, tabla referente a la comparación de dos operacionalizaciones

Poblaciones																	
Elemento principal	Sub-elemento	ID	Categoría	UPM01	UPM02	UPM03	UPM04	UPV05	US05	ORT05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	
Propiedades de los sujetos	Tipo de sujeto	5.1.1	Pregado avanzado	X	X	X	X	X	X	X	X	X	X	X	X	X	
		5.1.1	Grado (Máster o Doc.)													X	
		5.1.1	Profesional														X
	Experiencia en técnica (tratamiento)	5.1.2	Solo para experimento	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		5.1.2	La usa habitualmente (familiarizado)														
	Otras condiciones de los sujetos	5.1.3	Conocimientos de prog aptos p/ exp.	X	X	X	X	X	X			X	X	X	X	X	X
5.1.3		Conocimiento limitado de programación								X							
Popiedades de los objetos experimentales	Lenguaje	5.2.1	Lenguaje C	X	X	X	X	X	X	X	X	X	X	X	X	X	
	Tamaño	5.2.2	Pequeño (150-220 LOC)	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		5.2.2	Menos pequeño (221-300 LOC)	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Complejidad ciclomática	5.2.3	Media baja 15-29	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		5.2.3	Alta >45	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Tipos de Faltas	5.2.4	Grupo de Faltas G	X	X	X	X	X	X	X	X						
		5.2.4	Grupo de Faltas H	X	X	X	X	X	X	X	X						
		5.2.4	Grupo de Faltas I										X	X	X	X	X

Tabla G.27: **Partición 09**, tabla referente a las poblaciones

Estructura	Elemento principal	ID	Categoría	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04	UPV05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Experimentadores	Diseñador	1.1	Basili	X	X	X															
		1.1	Selby	X	X	X															
		1.1	Kamsties				X	X													
		1.1	Lott				X	X													
		1.1	Roper						X												
		1.1	Wood						X												
		1.1	Miller						X												
		1.1	Vegas							X	X	X	X	X	X	X	X				X
		1.1	Gómez																X	X	
		1.1	Solari													X	X	X			
		1.1	Apa																		X
	Entrenador	1.2	Basili	X	X	X															
		1.2	Selby	X	X	X															
		1.2	Kamsties				X	X													
		1.2	Lott				X	X													
		1.2	Roper						X												
		1.2	Wood						X												
		1.2	Miller						X												
		1.2	Juristo							X	X	X	X		X	X	X	X	X	X	X
		1.2	Vegas							X	X	X	X				X	X	X	X	X
		1.2	Letelier											X							
		1.2	Moreno												X	X					
	Monitor	1.3	Basili	X	X	X															
		1.3	Selby	X	X	X															
		1.3	Kamsties				X	X													
		1.3	Lott				X	X													
		1.3	Roper						X												
		1.3	Wood						X												
		1.3	Miller						X												
		1.3	Juristo							X	X	X	X				X	X	X	X	
		1.3	Vegas							X	X	X	X		X	X	X	X	X	X	X
		1.3	Dieste																		
		1.3	Ferre								X										
		1.3	Moreno												X	X					
		1.3	Gómez														X	X	X	X	
		1.3	Letelier											X							
		1.3	Apa																		X

Tabla G.28: **Partición 48**, tabla referente a los experimentadores y sitio

Estructura	Elemento principal	ID	Categoría	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04	UPV05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Experimentadores	Medición	1.4	Basili	X	X	X															
		1.4	Selby	X	X	X															
		1.4	Kamsties				X	X													
		1.4	Lott				X	X													
		1.4	Roper						X												
		1.4	Wood						X												
		1.4	Miller						X												
		1.4	Vegas							X	X	X	X							X	
		1.4	Gómez															X			
		1.4	Solari											X	X	X	X		X		
		1.4	Apa																		X
	Analista	1.5	Basili	X	X	X															
		1.5	Selby	X	X	X															
		1.5	Kamsties				X	X													
		1.5	Lott				X	X													
		1.5	Roper						X												
		1.5	Wood						X												
		1.5	Miller						X												
		1.5	Vegas							X	X	X	X	X	X						
		1.5	Gómez															X	X	X	
		1.5	Solari											X		X	X				
		1.5	Apa																		X
Sitio		2	UMD	X	X																
		2	NASA			X															
		2	UKL				X	X													
		2	UoS						X												
		2	UPM							X	X	X	X		X	X	X	X	X	X	X
		2	UPV											X							

Tabla G.29: **Partición 48**, tabla referente a los experimentadores y sitio (continuación)

		Protocolo experimental																			
Elemento principal	Sub-elemento	ID	Categoría	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04	UPV05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Análisis	Técnica de análisis de datos	3.1.1	ANOVA factorial							X	X	X	X	X	X	X	X	X	X	X	X
		3.1.1	ANOVA de una vía				X	X													
		3.1.1	ANOVA mixto	X	X	X			X												
Diseño	Diseño experimental	3.2.1	Diseño Factorial Mixto	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X
		3.2.1	Diseño Factorial																	X	X
		3.2.2	Todos los tipos de orden	X	X	X	X	X	X	X	X	X	X			X	X	X	X		
	Orden de aplicación de los tratamientos	3.2.2	E , F												X						
		3.2.2	No Aplica (c/ suj app 1 tec)																	X	X
		3.2.3	Intercalado												X						
	Combinación del entrenamiento con las sesiones	3.2.3	Secuencial	X	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X
		3.2.4	3 Técnicas, 1 Programa	X	X	X	X	X	X	X	X	X	X			X	X	X	X		
		3.2.4	1 Técnica, 3 Programas												X						
	Tratamientos por sesión	3.2.4	3 Técnicas, 3 Programas																	X	X
		3.2.4	2 Técnicas, 2 Progrmas																		X
		3.2.5	1 Sesión	X	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X
	Sesiones requeridas por sujeto para aplicar tratamientos	3.2.5	2 Sesiones												X						
		3.2.6	1 Técnica, 1 Programa	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		3.2.6	1 Técnica, 1 Programa	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Instrumentos	Instrumentos o equipo de medición	3.3.1	Ordenador, folios, bolígrafos (manual)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		3.3.1	Ordenador, folios, bolígrafos (manual)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Objetos experimentales	Programas	3.4.1	P1	X	X	X															
		3.4.1	P2	X	X																
		3.4.1	P3	X		X															
		3.4.1	P4		X	X															
		3.4.1	Cmdline				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		3.4.1	Nametbl				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		3.4.1	Ntree				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		3.4.1	Ntree				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		3.4.1	Ntree				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Faltas (Instancias)	3.4.2	Instancia A	X																	
		3.4.2	Instancia B		X																
		3.4.2	Instancia C			X															
		3.4.2	Instancia D				X														
		3.4.2	Instancia E					X													
		3.4.2	Instancia F						X												
Faltas (Instancias)	3.4.2	Instancia G							X	X	X	X	X	X	X						
	3.4.2	Instancia H							X	X	X	X	X	X	X						
	3.4.2	Instancia I														X	X	X	X	X	
	3.4.2	Instancia I														X	X	X	X	X	
	3.4.2	Instancia I														X	X	X	X	X	
	3.4.2	Instancia I														X	X	X	X	X	
Materiales	Formularios del experimento	3.5.1	Formularios v1	X	X	X															
		3.5.1	Formularios v2				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Instrucciones del experimento	3.5.2	Instrucciones v1	X	X	X															
		3.5.2	Instrucciones v2				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Guías suplementarias	3.5.3	Guías v1	X	X	X															
		3.5.3	Guías v2				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Tabla G.30: **Partición 48**, tabla referente al protocolo experimental

			Operacionalizaciones de los constructos																			
Elemento principal	Sub-elemento	ID	Categoría	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04	UPV05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11	
Causa	Tratamiento A	Tratamiento acorde a la literatura de SE		4.1.1.1.1	P.C.E. y A.V.L.	X	X	X	X	X	X											
				4.1.1.1.1	P.C.E.						X	X	X	X	X	X	X	X	X	X	X	
		Procedimiento de aplicación del tratamiento		4.1.1.1.2	Procedimiento de aplicación P.C.E. v1	X	X	X														
				4.1.1.1.2	Procedimiento de aplicación P.C.E. v2				X	X	X											
				4.1.1.1.2	Procedimiento de aplicación P.C.E. v3						X	X	X	X	X	X	X	X	X	X	X	X
		Tratamiento como instrucciones		4.1.1.1.3	P.C.E. v1	X	X	X														
				4.1.1.1.3	P.C.E. v2				X	X	X											
				4.1.1.1.3	P.C.E. v3							X	X	X	X	X	X					
				4.1.1.1.3	P.C.E. v4													X	X	X	X	X
		Recursos del tratamiento		4.1.1.1.4	Ordenador (1 sujeto x ordenador)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
				4.1.1.1.4	Impresora				X	X	X											
				4.1.1.1.4	Test Driver	X	X	X														
				4.1.1.1.4	Herramienta de cobertura de código	X	X	X	X	X	X											
		Aspectos de transmisión del tratamiento		4.1.1.1.5	Con experiencia en curso universitario	X	X		X	X	X	X	X	X			X	X	X	X	X	X
				4.1.1.1.5	Sin experiencia en curso universitario						X											
				4.1.1.1.5	Con experiencia en tutorial			X									X					
	Tratamiento B	Tratamiento acorde a la literatura de SE		4.1.2.1	Cobertura de sentencias	X	X	X														
				4.1.2.1	Cobertura de predicados				X	X												
				4.1.2.1	Cobertura de decisiones						X	X	X	X	X	X	X	X	X	X	X	
		Prodedimiento de aplicación del tratamiento		4.1.2.2	Procedimiento cobertura de sentencias	X	X	X														
				4.1.2.2	Procedimiento cobertura de predicados				X	X	X											
				4.1.2.2	Procedimiento cobertura de decisiones						X	X	X	X	X	X	X	X	X	X	X	
		Tratamiento como instrucciones		4.1.2.3	Cobertura de sentencias v1	X	X	X														
4.1.2.3				Cobertura de predicados v1				X	X													
4.1.2.3				Cobertura de decisiones v1						X												
4.1.2.3				Cobertura de decisiones v2							X	X	X	X	X							
4.1.2.3				Cobertura de decisiones v3												X	X	X	X			
4.1.2.3				Cobertura de decisiones v3																		
Recursos del tratamiento			4.1.2.4	Ordenador (1 persona x ord)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
			4.1.2.4	Impresora				X	X													
			4.1.2.4	Test Driver	X	X	X															
			4.1.2.4	Herramienta de cobertura de código	X	X	X	X	X	X												
Aspectos de transmisión del tratamiento			4.1.2.5	Con experiencia en curso universitario	X	X		X	X	X	X	X	X			X	X	X	X	X		
	4.1.2.5		Sin experiencia en curso universitario						X													
	4.1.2.5		Con experiencia en tutorial			X								X								
(Efectos)	Métrica		4.2.1	Porcentaje de f. observadas por sujeto	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X		
			4.2.1	Número de f. observadas por sujeto	X	X	X			X												
			4.2.2	Criterio relajado	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	
Efectividad	Procedimiento de medición (criterio de corrección)		4.2.2	Criterio estricto														X				

Tabla G.31: **Partición 48**, tabla referente a la comparación de dos operacionalizaciones

Poblaciones																					
Elemento principal	Sub-elemento	ID	Categoría	UMD82	UMD83	UMD84	UKL94	UKL95	UoS97	UPM01	UPM02	UPM03	UPM04	UPV05	UPM05	UPM06	UPM07	UPM08	UPM09	UPM10	UPM11
Propiedades de los sujetos	Tipo de sujeto	5.1.1	Pregrado avanzado	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	
		5.1.1	Grado (Máster o Doc.)	X	X															X	X
		5.1.1	Profesional			X															
	Experiencia en técnica (tratamiento)	5.1.2	Solo para experimento	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		5.1.2	La usa habitualmente (familiarizado)			X															
	Otras condiciones de los sujetos	5.1.3	Conocimientos de prog aptos p/ exp.	X	X	X			X	X	X	X	X	X	X	X	X	X	X	X	X
		5.1.3	Conocimiento limitado de programación				X	X													
Popiedades de los objetos	Lenguaje	5.2.1	Lenguaje Simplt	X	X																
		5.2.1	Lenguaje Fortran			X															
		5.2.1	Lenguaje C				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Tamaño	5.2.2	Más pequeño (< 150 LOC)	X	X	X															
		5.2.2	Pequeño (150-220 LOC)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		5.2.2	Menos pequeño (221-300 LOC)		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Complejidad ciclomática	5.2.3	Baja < 15	X		X															
		5.2.3	Media baja 15-29	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
		5.2.3	Media 30-45	X	X	X															
		5.2.3	Alta >45				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Tipos de Faltas	5.2.4	Grupo de Faltas A	X																	
		5.2.4	Grupo de Faltas B		X																
		5.2.4	Grupo de Faltas C			X															
		5.2.4	Grupo de Faltas D				X														
		5.2.4	Grupo de Faltas E					X													
		5.2.4	Grupo de Faltas F						X												
		5.2.4	Grupo de Faltas G							X	X	X	X	X	X	X					
		5.2.4	Grupo de Faltas H							X	X	X	X	X	X	X					
5.2.4		Grupo de Faltas I														X	X	X	X	X	

Tabla G.32: **Partición 48**, tabla referente a las poblaciones

Bibliografía

- ADAMS, A. T., AJROUCH, H., KRISTINE J. AND HENDERSON y IRENE, H. Service-learning outcomes research: The role and scarcity of replication studies. *Journal of Applied Sociology*, vol. 22(2), páginas 55–74, 2005.
- ALHAZEN. *Book of Optics*. 1021.
- ALMQVIST, J. P. F. *Replication of Controlled Experiments in Empirical Software Engineering – A Survey*. Proyecto Fin de Carrera, Department of Computer Science, Faculty of Science, Lund University, 2006. Supervisors: Amela Karahasanovic, Simula RL and Goran Fries, Lund University.
- ANDA, B. C., SJOBERG, D. I. y MOCKUS, A. Variability and reproducibility in software engineering: A study of four companies that developed the same system. *IEEE Transactions on Software Engineering*, vol. 35, páginas 407–429, 2009. ISSN 0098-5589.
- BAHR, H. M., CAPLOW, T. y CHADWICK, B. A. Middletown iii: Problems of replication, longitudinal measurement, and triangulation. *Annu. Rev. Sociol.*, vol. 9(1), páginas 243–264, 1983.
- BARKER, H. R. y GURMAN, E. B. Replication versus tests of equivalence. *Perceptual and Motor Skills*, vol. 35, páginas 807–814, 1972.
- BASIL, V. y SELBY, R. Comparing the effectiveness of software testing strategies. *IEEE Trans. Softw. Eng.*, vol. 13(12), páginas 1278–1296, 1987. ISSN 0098-5589.
- BASIL, V., SELBY, R. y HUTCHENS, D. Experimentation in software engineering. *IEEE Trans. Softw. Eng.*, vol. 12(7), páginas 733–743, 1986.
- BASIL, V., SHULL, F. y LANUBILE, F. Building knowledge through families of experiments. *IEEE Transactions on Software Engineering*, vol. 25(4), páginas 456–473, 1999.
- BEAULIEU, A. Transversely excited atmospheric pressure co2 lasers. *Applied Physics Letters*, vol. 16(12), páginas 504–505, 1970.
- BECK, C. T. Replication strategies for nursing research. *Journal of Nursing Scholarship*, vol. 26, páginas 191–194, 1994. ISSN 1547-5069.
- BELOFF, J. Research strategies for dealing with unstable phenomena. En *The repeatability problem in parapsychology* (editado por B. Shapin y L. Coly). The Parapsychology Foundation, 1985.
- BISWAS, D. J. y HARRISON, R. G. Observation of pulsating instabilities and chaos in a transversely excited atmospheric pressure co2 laser. *Optics Communications*, vol. 57(3), páginas 193–195, 1986. ISSN 0030-4018.

- BLOMQUIST, K. B. Replication of research. *Res. Nurs. Health*, vol. 9(3), páginas 193–194, 1986. ISSN 1098-240X.
- BORENSTEIN, M., HEDGES, L. V., HIGGINS, J. P. y ROTHSTEIN, H. R. *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd, United Kingdom, 2009.
- BRINBERG, D. y MCGRATH, J. E. *Validity and the Research Process*. Sage Publications, Inc., Newbury Park, Calif., 1985. ISBN 0803923021 0803933762.
- BROOKS, A., DALY, J., MILLER, J., ROPER, M. y WOOD, M. Replication of experimental results in software engineering. Informe Técnico ISERN-96-10, Scotland, 1996.
- BROWN, S. W. y CONEY, K. A. Building a replication tradition in marketing. En *Marketing 1776-1976 and beyond* (editado por K. Bernhardt). Chicago: American Marketing Association, 1976.
- BROWNE, M. Physicists debunk claim of a new kind of fusion. *The New York Times*, 1989.
- CAILLIEZ, F. The analytical solution of the additive constant problem. *Psychometrika*, vol. 48, páginas 305–308, 1983. ISSN 0033-3123.
- CAMPBELL, D. T. y STANLEY, J. C. *Experimental and Quasi-Experimental Designs for Research*. Houghton Mifflin Company, 1963.
- CARTWRIGHT, N. Replicability, reproducibility, and robustness: Comments on harry collins. *History of Political Economy*, vol. 23(1), páginas 143–155, 1991.
- CHESSEL, D., DUFOUR, A. y THIOULOUSE, J. The ade4 package-I- One-table methods. *R News*, vol. 4, páginas 5–10, 2004.
- COCHRAN, W. G. Problems arising in the analysis of a series of similar experiments. *Journal of the Royal Statistical Society*, vol. 4(Suppl.), páginas 102–118, 1937.
- COCHRAN, W. G. The combination of estimates from different experiments. *Biometrics*, vol. 10(1), páginas 101–129, 1954. ISSN 0006341X.
- COHEN, J. *Statistical Power Analysis for the Behavioral Sciences*. L. Erlbaum Associates, Hillsdale, NJ, 1988.
- COLLINS, H. M. *Changing order: Replication and Induction in Scientific Practice*. Sage Publications, 1985.
- COOK, T. y CAMPBELL, D. *The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings*. Rand McNally, Chicago, 1976.
- CROSON, R. The method of experimental economics. *International Negotiation*, vol. 10, páginas 131–148(18), 2005.
- CRUZ-LEMUS, J. A., GENERO, M., MANSO, M. E. y PIATTINI, M. Evaluating the effect of composite states on the understandability of uml statechart diagrams. En *8th International Conference on Model-Driven Engineering, Languages and Systems (MoDELS 2005)*, páginas 113–125. 2005.
- DIESTE, O., GRIMÁN, A. y JURISTO, N. Developing search strategies for detecting relevant experiments. *Empirical Softw. Engg.*, vol. 14, páginas 513–539, 2009. ISSN 1382-3256.
- DRAY, S. y DUFOUR, A. The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, vol. 22(4), páginas 1–20, 2007.

- DRAY, S., DUFOUR, A. y CHESSEL, D. The ade4 package-II: Two-table and K-table methods. *R News*, vol. 7(2), páginas 47–52, 2007.
- DYBÅ, T., KAMPENES, V. y SJØBERG, D. A systematic review of statistical power in software engineering experiments. *Information and Software Technology*, vol. 48(8), páginas 745–755, 2006.
- EASLEY, R., MADDEN, C. y DUNN, M. Conducting marketing science: The role of replication in the research process. *Journal of Business Research*, vol. 48(1), páginas 83–92, 2000.
- EASTERBROOK, S., SINGER, J., STOREY, M. y DAMIAN, D. Selecting empirical methods for software engineering research. En *Guide to Advanced Empirical Software Engineering* (editado por F. Shull, J. Singer y D. I. Sjøberg). Springer, London, 2008.
- ERIKSSON, H.-E. y PENKER, M. *Business Modeling With UML: Business Patterns at Work*. John Wiley & Sons, Inc., New York, NY, USA, 1998. ISBN 0471295515.
- EVANSCHITZKY, H. y ARMSTRONG, J. S. Replications of forecasting research. *International Journal of Forecasting*, vol. 26(1), páginas 4–8, 2010. ISSN 0169-2070. Special Section: European Election Forecasting.
- EVERITT, B. S., LANDAU, S. y LEESE, M. *Cluster Analysis*. Arnold Publishers, 2001. ISBN 0340761199.
- FINIFTER, B. The generation of confidence: Evaluating research findings by random subsample replication. *Sociological Methodology*, vol. 4, páginas 112–175, 1972.
- FINIFTER, B. Replication and extension of social research through secondary analysis. *Social Science Information/Information sur les Sciences Sociales*, vol. 14(2), páginas 119–153, 1975. ISSN 05390184.
- FISHER, R. A. *Statistical Methods for Research Workers*. Oliver & Boyd, London, 4th ed. edición, 1932.
- FLEISCHMANN, M. y PONS, S. Electrochemically induced nuclear fusion of deuterium. *Journal of Electroanalytical Chemistry*, vol. 261(2, Part 1), páginas 301 – 308, 1989. ISSN 0022-0728.
- FOMEL, S. y CLAERBOUT, J. F. Guest editors' introduction: Reproducible research. *Computing in Science and Engineering*, vol. 11, páginas 5–7, 2009. ISSN 1521-9615.
- FRIEDMAN, D. y SUNDER, S. *Experimental Methods: A Primer for Economists*. Cambridge University Press, 1994.
- FUESS, S. M. On replications in business and economics research: the qjbe case. *Quarterly Journal of Business and Economics*, 1996.
- FUSARO, P., LANUBILE, F. y VISAGGIO, G. A replicated experiment to assess requirements inspection techniques. *Empirical Softw. Engg.*, vol. 2(1), páginas 39–57, 1997. ISSN 1382-3256.
- GEORGE, E. O. *Combining Independent One-sided and Two-sided Statistical tests – Some Theory and Applications*. Tesis Doctoral, University of Rochester, 1977.
- GIBSON, A. F., KIMMITT, M. F. y PATEL, B. S. The gain of transversely excited co2 lasers at pressures up to 500 torr. *J. Phys. D: Appl. Phys.*, vol. 4(7), páginas 882–887, 1971.
- GILES, J. The trouble with replication. *Nature*, vol. 442(7101), páginas 344–347, 2006.
- GLASS, G. V. Primary, secondary, and meta-analysis of research. *Educational Researcher*, vol. 5(10), páginas 3–8, 1976.

- GOWER, J. C. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, vol. 53, páginas 325–338, 1966.
- GOWER, J. C. y LEGENDRE, P. Metric and euclidean properties of dissimilarity coefficients. *Journal of Classification*, vol. 3, páginas 5–48, 1986.
- GRANT, E. E. y SACKMAN, H. An exploratory investigation of programmer performance under on-line and off-line conditions. *IEEE Transactions on Human Factors in Electronics*, vol. 8(1), páginas 33–48, 1967.
- GUALA, F. *The Methodology of Experimental Economics*. Cambridge University Press, 2005.
- GÓMEZ, O. S., JURISTO, N. y VEGAS, S. Replication, reproduction and re-analysis: Three ways for verifying experimental findings. En *International Workshop on Replication in Empirical Software Engineering Research (RESER'2010)*. Cape Town, South Africa, 2010a.
- GÓMEZ, O. S., JURISTO, N. y VEGAS, S. Replications types in experimental disciplines. En *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM'2010)*, ESEM '10, páginas 3:1–3:10. ACM, Bolzano, Italy, 2010b. ISBN 978-1-4503-0039-1.
- HANNAY, J., DYBÅ, T., ARISHOLM, E. y SJØBERG, D. The effectiveness of pair programming: A meta-analysis. *Information and Software Technology*, vol. 51(7), páginas 1110–1122, 2009.
- HANNAY, J., SJØBERG, D. y DYBÅ, T. A systematic review of theory use in software engineering experiments. *Software Engineering, IEEE Transactions on*, vol. 33(2), páginas 87–107, 2007. ISSN 0098-5589.
- HANSON, D. R. Ratsno - an experiment in software adaptability. *Software - Practice and Experience*, vol. 7(5), páginas 625–630, 1977.
- HARTLEY, H. O. The range in random samples. *Biometrika*, vol. 32, páginas 334–348, 1942.
- HARTLEY, H. O. y PEARSON, E. S. The probability integral of the range in samples of n observations from a normal population. *Biometrika*, vol. 32, páginas 301–310, 1942.
- HAYES, W. Research synthesis in software engineering: A case for meta-analysis. En *METRICS '99: Proceedings of the 6th International Symposium on Software Metrics*, página 143. IEEE Computer Society, Washington, DC, USA, 1999.
- HCIL, U. o. M. Treemap. <http://www.cs.umd.edu/hcil/treemap/>, 2004.
- HEDGES, L. V. y OLKIN, I. *Statistical Methods for Meta-analysis*. Academic Press, Orlando :, 1985. ISBN 0123363802 0123363802.
- HENDRICK, C. *Replications, Strict Replications, and Conceptual Replications: Are They Important?*. Sage, Newbury Park, California, 1990.
- HIGGINS, J. P. T. y THOMPSON, S. G. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, vol. 21(11), páginas 1539–1558, 2002.
- HOWDEN, W. Functional program testing. *IEEE Transactions on Software Engineering*, vol. 6, páginas 162–169, 1980. ISSN 0098-5589.
- HUEDO-MEDINA, T. B., SÁNCHEZ-MECA, J., MARÍN-MARTÍNEZ, F. y BOTELLA, J. Assessing heterogeneity in meta-analysis: q statistic or i^2 index? *Psychological Methods*, vol. 11(2), páginas 193–206, 2006.
- HULL, L. W. H. *Historia y Filosofía de la Ciencia*. Crítica, Barcelona, 2011.

- HUNT, K. Do we really need more replications? *Psychological Reports*, vol. 36(2), páginas 587–593, 1975. ISSN 0033-2941.
- HUNTER, J. The desperate need for replications. *Journal of Consumer Research*, vol. 28(1), páginas 149–158, 2001.
- HUNTER, J. E. y SCHMIDT, F. L. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Sage Publications, Inc, 2nd ed. edición, 2004. ISBN 1412909120.
- HYMAN, H. H. y WRIGHT, C. R. Evaluating social action programs. En *The Uses of Sociology* (editado por H. L. Lazarsfeld y H. L. Wilensky), páginas 769–777. NY Basic, 1967.
- INFOMATA. Datafile. <http://datafile.sourceforge.net/index.shtml>, 2009.
- ISHIKAWA, K. *Introduction to Quality Control*. Productivity Press, New York, 1 edición, 1990. ISBN 490622461X.
- ISO. Iso 5725-1, accuracy (trueness and precision) of measurement methods and results. 1994.
- ISO. Iso/iec 9126, software product evaluation – quality characteristics and guidelines for their use. 1999.
- JACCARD, P. Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudense des Sciences Naturelles*, vol. 44, páginas 223–270, 1908.
- JEDLITSCHKA, A. y PFAHL, D. Reporting guidelines for controlled experiments in software engineering. *ACM/IEEE International Symposium on Empirical Software Engineering*, 2005.
- JIANG, Y., JAHAGIRDAR, B. N., REINHARDT, R. L., SCHWARTZ, R. E., KEENE, C. D., ORTIZ-GONZALEZ, X. R., REYES, M., LENVIK, T., LUND, T., BLACKSTAD, M., DU, J., ALDRICH, S., LISBERG, A., LOW, W. C., LARGAESPADA, D. A. y VERFAILLIE, C. M. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature*, vol. 418(6893), páginas 41–49, 2002. ISSN 0028-0836.
- JONES, L. V. y FISKE, D. W. Models for testing the significance of combined results. *Psychological Bulletin*, vol. 50(5), páginas 375–382, 1953. ISSN 0033-2909.
- JURISTO, N. y GÓMEZ, O. S. Replication of software engineering experiments. En *Empirical Software Engineering and Verification: LASER Summer School 2008-2010* (editado por B. Meyer y M. Nordio), vol. 7007 de *Lecture Notes in Computer Science*, páginas 60–88. Springer-Verlag, Berlin, 2011. ISBN 978-3-642-25230-3.
- JURISTO, N., MORENO, A. y VEGAS, S. Reviewing 25 years of testing technique experiments. *Empirical Softw. Engg.*, vol. 9(1-2), páginas 7–44, 2004. ISSN 1382-3256.
- JURISTO, N. y MORENO, A. M. *Basics of Software Engineering Experimentation*. Kluwer Academic Publishers, 2001.
- JURISTO, N., MORENO, A. M. y VEGAS, S. A survey on testing technique empirical studies: How limited is our knowledge. En *Proceedings of the 2002 International Symposium on Empirical Software Engineering*, páginas 161–162. IEEE Computer Society, Washington, DC, USA, 2002. ISBN 0-7695-1796-X.
- JURISTO, N. y VEGAS, S. Functional testing, structural testing and code reading: What fault type do they each detect? En *Empirical Methods and Studies in Software Engineering* (editado por R. Conradi y A. Wang), vol. 2765 de *Lecture Notes in Computer Science*, páginas 208–232. Springer Berlin / Heidelberg, 2003.

- JURISTO, N. y VEGAS, S. Using differences among replications of software engineering experiments to gain knowledge. En *ESEM '09: Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement*, páginas 356–366. IEEE Computer Society, Washington, DC, USA, 2009.
- JØRGENSEN, M. A review of studies on expert estimation of software development effort. *Journal of Systems and Software*, vol. 70(1-2), páginas 37–60, 2004.
- KAMSTIES, E. y LOTT, C. An empirical evaluation of three defect detection techniques. Informe Técnico ISERN 95-02, Dept. Computer Science, University of Kaiserslautern, Kaiserslautern, Germany, 1995a.
- KAMSTIES, E. y LOTT, C. M. An empirical evaluation of three defect-detection techniques. En *Proceedings of the 5th European Software Engineering Conference*, páginas 362–383. Springer-Verlag, London, UK, 1995b.
- KANTOWITZ, B. H., ROEDIGER III, H. L. y ELMES, D. G. *Experimental Psychology*. Wadsworth Publishing, second ed. edición, 1984.
- KELLY, C., CHASE, L. y TUCKER, R. Replication in experimental communication research: an analysis. *Human Communication Research*, vol. 5(4), páginas 338–342, 1979.
- KEPPEL, G. *Design and Analysis – A Researcher's Handbook*. Prentice Hall, Englewood Cliffs, NJ., 1982.
- KITCHENHAM, B. The role of replications in empirical software engineering—a word of warning. *Empirical Softw. Engg.*, vol. 13(2), páginas 219–221, 2008.
- KITCHENHAM, B., PFLEGER, S., PICKARD, L., JONES, P., HOAGLIN, D., EL EMAM, K. y ROSENBERG, J. Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering*, vol. 28(8), páginas 721–734, 2002.
- KOLENCE, K. W. Experiments & measures in computing. páginas 69–72, 1973.
- KOVACEVIC, J. How to encourage and publish reproducible research. En *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP '2007)*, vol. 4, páginas IV–1273–IV–1276. 2007.
- KREIN, J. L. y KNUTSON, C. D. A case for replication: Synthesizing research methodologies in software engineering. En *1st International Workshop on Replication in Empirical Software Engineering Research (RESER '2010)*. Cape Town, South Africa, 2010.
- LA SORTE, M. A. Replication as a verification technique in survey research: A paradigm. *The Sociological Quarterly*, vol. 13(2), páginas 218–227, 1972.
- LEONE, R. y SCHULTZ, R. A study of marketing generalizations. *The Journal of Marketing*, vol. 44(1), páginas 10–18, 1980.
- LINDSAY, R. M. y EHRENBERG, A. S. C. The design of replicated studies. *The American Statistician*, vol. 47(3), páginas 217–228, 1993.
- LINGER, R. C., WITT, B. I. y MILLS, H. D. *Structured Programming; Theory and Practice the Systems Programming Series*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1979. ISBN 0201144611.
- LINGOES, J. Some boundary conditions for a monotone analysis of symmetric matrices. *Psychometrika*, vol. 36, páginas 195–203, 1971. ISSN 0033-3123.
- LOTT, C. y ROMBACH, H. D. Repeatable software engineering experiments for comparing defect-detection techniques. *Empirical Software Engineering*, vol. 1, páginas 241–277, 1996.

- LUNG, J., ARANDA, J., EASTERBROOK, S. y WILSON, G. On the difficulty of replicating human subjects studies in software engineering. En *ICSE '08: Proceedings of the 30th international conference on Software engineering*, páginas 191–200. ACM, New York, NY, USA, 2008. ISBN 978-1-60558-079-1.
- LYKKEN, D. T. Statistical significance in psychological research. *Psychol Bull*, vol. 70(3), páginas 151–159, 1968.
- LÜDTKE, R. Do that to me one more time! – what kind of trial replications do we need? *Complementary Therapies in Medicine*, vol. 16(4), páginas 181–2, 2008.
- MANDIĆ, V., MARKKULA, J. y OIVO, M. Towards multi-method research approach in empirical software engineering. En *Lecture Notes in Business Information Processing* (editado por F. Bomarius, M. Oivo, P. Jaring y P. Abrahamsson), vol. 32, páginas 96–110. Springer Berlin Heidelberg, 2009.
- MARICK, B. *The Craft of Software Testing: Subsystem Testing Including Object-based and Object-Oriented Testing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1995. ISBN 0-13-177411-5.
- MENDONÇA, M., MALDONADO, J., DE OLIVEIRA, M., CARVER, J., FABBRI, S., SHULL, F., TRAVASSOS, G., HÖHN, E. y BASILI, V. A framework for software engineering experimental replications. En *Proceedings of the 13th IEEE International Conference on Engineering of Complex Computer Systems (ICECCS '08)*, páginas 203–212. IEEE Computer Society, Washington, DC, USA, 2008.
- MEYER, B. The rise of empirical software engineering (ii): What we are still missing. Blog, 2009.
- MILLER, J. Can results from software engineering experiments be safely combined? En *METRICS '99: Proceedings of the 6th International Symposium on Software Metrics*, página 152. IEEE Computer Society, Washington, DC, USA, 1999.
- MILLER, J. Applying meta-analytical procedures to software engineering experiments. *J. Syst. Softw.*, vol. 54(1), páginas 29–39, 2000.
- MILLER, J. Replicating software engineering experiments: A poisoned chalice or the holy grail. *Information and Software Technology*, vol. 47(4), páginas 233–244, 2005.
- MILLER, J. Triangulation as a basis for knowledge discovery in software engineering. *Empirical Softw. Engg.*, vol. 13(2), páginas 223–228, 2008.
- MILLER, J., WOOD, M. y ROPER, M. Further experiences with scenarios and checklists. *Empirical Software Engineering*, vol. 3(1), páginas 37–64, 1998.
- MITTELSTAEDT, R. y ZORN, T. Econometric replication: Lessons from the experimental sciences. *Quarterly Journal of Business & Economics*, vol. 23(1), 1984.
- MONROE, K. B. Front matter. *The Journal of Consumer Research*, vol. 19(1), páginas i–iv, 1992.
- MONTGOMERY, D. C. *Introduction to Statistical Quality Control*. John Wiley & Sons, 1996.
- MYERS, G. J. *The Art of Software Testing*. John Wiley & Sons, New York, 1979.
- NELSON, L. S. Technical aids. *Journal of Quality Technology*, vol. 16(4), páginas 238–239, 1984.
- OAKLAND, J. S. *Statistical Process Control*. Butterworth-Heinemann, 5 ed. edición, 2003.
- PARK, C. L. What is the value of replicating other studies? *Research Evaluation*, vol. 13, páginas 189–195(7), 2004.

- PATEL, B. S. Transversely excited high pressure carbon-dioxide laser. *Proc. Indian Nat. Sci. Acad., Part A*, vol. 39(5), páginas 309–313, 1973.
- PEARSON, K. Report on certain enteric fever inoculation statistics. *Br. Med. J.*, vol. 5(2), páginas 1243–1246, 1904.
- PERRY, D., PORTER, A. y VOTTA, L. Empirical studies of software engineering: A roadmap. En *ICSE '00: Proceedings of the Conference on The Future of Software Engineering*, páginas 345–355. ACM, New York, NY, USA, 2000.
- PFLEEGER, S. L. Albert einstein and empirical software engineering. *Computer*, vol. 32(10), páginas 32–38, 1999. ISSN 0018-9162.
- PFLEEGER, S. L. Soup or art? the role of evidential force in empirical software engineering. *IEEE Softw.*, vol. 22(1), páginas 66–73, 2005. ISSN 0740-7459.
- PICKARD, L., KITCHENHAM, B. y JONES, P. Combining empirical results in software engineering. *Information and Software Technology*, vol. 40(14), páginas 811–821, 1998.
- PORTER, A. y VOTTA, L. Comparing detection methods for software requirements inspections: A replication using professional subjects. *Empirical Software Engineering*, vol. 3(4), páginas 355–379, 1998.
- PORTER, A., VOTTA, L. y BASILI, V. Comparing detection methods for software requirements inspections: A replicated experiment. *IEEE Trans. Softw. Eng.*, vol. 21(6), páginas 563–575, 1995. ISSN 0098-5589.
- R DEV. CORE TEAM. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011.
- RADDER, H. Experimental reproducibility and the experimenters' regress. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 1, páginas 63–73, 1992.
- RADDER, H. *In and About the World: Philosophical Studies of Science and Technology*. State University of New York Press, Albany, N.Y., 1996.
- RAO, K. On the question of replication. *Journal of Parapsychology*, vol. 45, páginas 311–320, 1981.
- REID, L. N., SOLEY, L. C. y WIMMER, R. D. Replication in advertising research: 1977, 1978, 1979. *Journal of Advertising*, vol. 10(1), páginas 3–13, 1981.
- ROGERS, D. J. y TANIMOTO, T. T. A computer program for classifying plants. *Science*, vol. 132(3434), páginas 1115–1118, 1960.
- ROPER, M., WOOD, M. y MILLER, J. An empirical evaluation of defect detection techniques. *Information and Software Technology*, vol. 39(11), páginas 763–775, 1997.
- ROSENTHAL, R. Combining results of independent studies. *Psychological Bulletin*, vol. 85(1), páginas 185–193, 1978. ISSN 0033-2909.
- ROSENTHAL, R. Replication in behavioral research. *Journal of Social Behavior and Personality*, vol. 4(4), páginas 1–30, 1990.
- RUNESON, P., ANDERSSON, C., THELIN, T., ANDREWS, A. y BERLING, T. What do we know about defect detection methods? *IEEE Softw.*, vol. 23(3), páginas 82–90, 2006. ISSN 0740-7459.

- SANDAHL, K., BLOMKVIST, O., KARLSSON, J., KRYSANDER, C., LINDVALL, M. y OHLSSON, N. An extended replication of an experiment for assessing methods for software requirements inspections. *Empirical Software Engineering*, vol. 3(4), páginas 327–354, 1998.
- SARGENT, C. L. The repeatability of significance and the significance of repeatability. *European Journal of Parapsychology*, vol. 3, páginas 423–433, 1981.
- SCHMIDT, S. Shall we really do it again? the powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, vol. 13(2), páginas 90–100, 2009.
- SCHWARZER, G. *meta: Meta-Analysis with R*, 2010. R package version 1.6-1.
- SCRUCCA, L. qcc: an r package for quality control charting and statistical process control. *R News*, vol. 4/1, páginas 11–17, 2004.
- SELBY, R. Empirical software engineering research roadmap: Discussion and summary. En *Empirical Software Engineering Issues. Critical Assessment and Future Directions* (editado por V. Basili, D. Rombach, K. Schneider, B. Kitchenham, D. Pfahl y R. Selby), vol. 4336 de *Lecture Notes in Computer Science*, páginas 184–187. Springer Berlin / Heidelberg, 2007.
- SHAW, M. Prospects for an engineering discipline of software. *IEEE Softw.*, vol. 7(6), páginas 15–24, 1990. ISSN 0740-7459.
- SHAW, M. Continuing prospects for an engineering discipline of software. *IEEE Software*, vol. 26(6), páginas 64–67, 2009. ISSN 0740-7459.
- SHEWHART, W. A. *Economic Control of Quality of Manufactured Product*. Van Nostrand Reinhold Co., New York, 1931.
- SHNEIDERMAN, B. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Trans. Graph.*, vol. 11, páginas 92–99, 1992. ISSN 0730-0301.
- SHULL, F., BASILI, V., CARVER, J., MALDONADO, J., TRAVASSOS, G., MENDONÇA, M. y FABRI, S. Replicating software engineering experiments: Addressing the tacit knowledge problem. En *ISESE '02: Proceedings of the 2002 International Symposium on Empirical Software Engineering*, página 7. IEEE Computer Society, Washington, DC, USA, 2002.
- SHULL, F., CARVER, J., VEGAS, S. y JURISTO, N. The role of replications in empirical software engineering. *Empirical Softw. Engg.*, vol. 13(2), páginas 211–218, 2008.
- SHULL, F. y FELDMANN, R. L. Building theories from multiple evidence sources. *Guide to Advanced Empirical Software Engineering*, páginas 337–364, 2008.
- SIDMAN, M. *Tactics of Scientific Research*. NY Basic, 1960.
- SINGER, J. Using the american psychological association (apa) style guidelines to report experimental results. En *Proceedings of the Fifth IEEE Workshop on Empirical Studies of Software Maintenance (WESS'99)*, páginas 71–75. 1999.
- SJØBERG, D., DYBÅ, T. y JORGENSEN, M. The future of empirical methods in software engineering research. *Future of Software Engineering, (FOSE '07)*, páginas 358–378, 2007.
- SJØBERG, D., HANNAY, J., HANSEN, O., KAMPENES, V., KARAHASANOVIC, A., LIBORG, N.-K. y REKDAL, A. A survey of controlled experiments in software engineering. *Software Engineering, IEEE Transactions on*, vol. 31(9), páginas 733–753, 2005.
- SJØBERG, D. I. K., DYBÅ, T., ANDA, B. C. D. y HANNAY, J. E. Building theories in software engineering. *Guide to Advanced Empirical Software Engineering*, páginas 312–336, 2008.

- SMITH, V. L. Experimental economics: Induced value theory. *American Economic Review*, vol. 66(2), páginas 274–79, 1976.
- SMITH, V. L. Microeconomic systems as an experimental science. *American Economic Review*, vol. 72(5), páginas 923–55, 1982.
- SNEATH, P. The application of computers to taxonomy. *Journal of general microbiology*, vol. 17, páginas 201–226, 1957.
- SOKAL, R. R. y MICHENER, C. D. A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin*, vol. 28, páginas 1409–1438, 1958.
- SOKAL, R. R. y SNEATH, P. H. A. *Principles of Numerical Taxonomy*. W. H. Freeman and Company, San Francisco, 1963.
- SØRENSEN, T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. *Biol. Skr.*, vol. 5, páginas 1–34, 1948.
- STOUFFER, S. A., SUCHMAN, E. A., DEVINNEY, L. C., STAR, S. A. y WILLIAMS, R. M., JR. *The American soldier*, vol. I. Princeton University Press, Princeton, NJ, 1949.
- TICHY, W. Should computer scientists experiment more? *Computer*, vol. 31(5), páginas 32–40, 1998.
- TIPPETT, L. H. C. *The Method of Statistics*. Williams & Norgate, London, 1931.
- TROCHIM, W. *The Research Methods Knowledge Base, Second Edition*. Atomic Dog Publishing, Cincinnati, OH, 2nd edición, 2000. ISBN 1931442487.
- TSANG, E. y KWAN, K.-M. Replication and theory development in organizational science: A critical realist perspective. *The Academy of Management Review*, vol. 24(4), páginas 759–780, 1999.
- VAN IJZENDOORN, M. H. A process model of replication studies: On the relation between different types of replication. *Leiden University Library*, 1994.
- VEGAS, S., JURISTO, N., MORENO, A., SOLARI, M. y LETELIER, P. Analysis of the influence of communication between researchers on experiment replication. En *ISESE '06: Proceedings of the 2006 ACM/IEEE international symposium on Empirical software engineering*, páginas 28–37. ACM, New York, NY, USA, 2006.
- WASSERMAN, J. D. y BRACKEN, B. A. Psychometric characteristics of assessment procedures. En *Handbook of psychology: Assessment psychology, Vol. 10*. (editado por J. R. Graham y J. A. Naglieri), páginas 43–66. Hoboken, NJ, US: John Wiley & Sons Inc, 2003.
- WECO. *Western Electric Statistical Quality Control Handbook*. Western Electric Co, Indianapolis, IN, USA, 1956.
- WEINBERG, G. M. y GRESSETT, G. L. An experiment in automatic verification of programs. *Commun. ACM*, vol. 6, páginas 610–613, 1963. ISSN 0001-0782.
- WHEELER, D. J. *Understanding Variation: The Key to Managing Chaos*. SPC Press, Inc., Knoxville, Tennessee, 1993. ISBN 0945320531.
- WIKIPEDIA. Reproducibility — wikipedia, the free encyclopedia. 2009. [Online; accessed 9-September-2009].
- WIKIPEDIA. Alhazen — wikipedia, the free encyclopedia. 2011. [Online; accessed 8-August-2011].

- WILLIAMS, L., MCDOWELL, C., NAGAPPAN, N., FERNALD, J. y WERNER, L. Building pair programming knowledge through a family of experiments. En *Proceedings of the 2003 International Symposium on Empirical Software Engineering*, ISESE '03, páginas 143–. IEEE Computer Society, Washington, DC, USA, 2003. ISBN 0-7695-2002-2.
- WINER, B. J. *Statistical Principles in Experimental Design*. McGraw-Hill, New York, 2nd ed. edición, 1971. ISBN 0070709815.
- WOHLIN, C., RUNESON, P., HÖST, M., OHLSSON, M. C., REGNELL, B. y WESSLÉN, A. *Experimentation in Software Engineering: An Introduction*. Kluwer Academic Publishers, Boston, 1999.
- WOOD, M., ROPER, M., BROOKS, A. y MILLER, J. Comparing and combining software defect detection techniques: A replicated empirical study. *SIGSOFT Softw. Eng. Notes*, vol. 22(6), páginas 262–277, 1997. ISSN 0163-5948.
- ZANNIER, C., MELNIK, G. y MAURER, F. On the success of empirical studies in the international conference on software engineering. En *ICSE '06: Proceedings of the 28th international conference on Software engineering*, páginas 341–350. ACM, New York, NY, USA, 2006.
- ZELKOWITZ, M. V., WALLACE, D. R. y BINKLEY, D. W. *Experimental Validation of new Software Technology*, páginas 229–263. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2003. ISBN 981-02-4914-4.

Lista de acrónimos

AHC.....	<i>Agglomerative Hierarchical Clustering</i>
ANOVA.....	<i>Analysis of Variance</i> , Análisis de la varianza
APA.....	<i>American Psychological Association</i>
AVL.....	Análisis de Valores Límite
BVA.....	<i>Boundary Value Analysis</i> , Análisis de Valores Límite
CONACYT...	Consejo Nacional de Ciencia y Tecnología
CSV.....	<i>Comma-separated values</i>
EPS.....	<i>Encapsulated PostScript</i>
ESE.....	Experimental Software Engineering
FI-UPM.....	Facultad de Informática de la Universidad Politécnica de Madrid
GRISE.....	Grupo de Investigación en Ingeniería del Software Empírica
HCIL.....	<i>Human-Computer Interaction Laboratory</i> , Laboratorio de Interacción Persona-Ordenador
ICSE.....	<i>International Conference on Software Engineering</i>
IS.....	Ingeniería del Software
ISE.....	Ingeniería del Software Experimental
JAR.....	<i>Java Archive</i>
JDK.....	<i>Java Development Kit</i>
LCL.....	<i>Lower Control Limit</i> , Límite de Control Inferior
LWL.....	<i>Lower Warning Limit</i> , Límite de Alerta Inferior
MDS.....	<i>Multidimensional Scaling</i> , Escalado Multidimensional
PCE.....	Particiones de Clases de Equivalencia
SE.....	Software Engineering
SPC.....	<i>Statistical Process Control</i>
UCL.....	<i>Upper Control Limit</i> , Límite de Control Superior
UML.....	<i>Unified Modeling Language</i> , Lenguaje de Modelado Unificado
UWL.....	<i>Upper Warning Limit</i> , Límite de Alerta Superior
WEC.....	<i>Western Electric Company</i>
WMD.....	<i>Weighted Mean Difference</i>
WSS.....	<i>Weighted Sum of Squares</i>

*Nube de palabras de esta tesis formada por
250 palabras repetidas al menos 60 veces.*





POLITÉCNICA